Unit C4

Eigenvectors

Introduction

By now you should be familiar with a wide variety of linear transformations from one vector space to another, and should appreciate that the matrix of a linear transformation depends on the bases chosen for the domain and codomain. In this final unit on linear algebra we concentrate on linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$, from $\mathbb{R}^3$ to $\mathbb{R}^3$ and, more generally, from $\mathbb{R}^n$ to $\mathbb{R}^n$, and address the following question.

> Is it possible to find a basis for both the domain and codomain so that the matrix of a linear transformation is a diagonal matrix?

In the preceding units of this book you have studied vectors, matrices, vector spaces and linear transformations. The method for finding a diagonal matrix of a linear transformation (if such a matrix exists) links all these topics together. To round off the linear algebra topic, we use linear transformations and diagonal matrices to classify conics and quadrics.

1 Eigenvalues and eigenvectors

In this section you will see that some lines through the origin are mapped to themselves by some linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$: the individual points on these lines are usually moved, but, for a given line, all the points are scaled by a constant factor. You will see that this idea of fixed lines also applies to linear transformations from $\mathbb{R}^3$ to $\mathbb{R}^3$ and, more generally, from $\mathbb{R}^n$ to $\mathbb{R}^n$. You will learn how determinants can be used for finding these fixed lines of linear transformations.

1.1 What is an eigenvector?

In Subsection 1.1 of Unit C3 *Linear transformations* you saw that a linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ moves the points of the plane around, but fixes the origin. Furthermore, parallel lines get mapped to parallel lines. In this section we will observe that t may map some lines through the origin onto themselves. These 'unchanged' lines are rather special.

Consider the linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$t(x, y) = (x + 4y, x - 2y).$$

We know that t maps the origin $(0, 0)$ to itself, since this is a property of all linear transformations.

We can calculate the image of the point $(1, 0)$:

$$t(1, 0) = (1 + (4 \times 0), 1 - (2 \times 0)) = (1, 1).$$

Since linear transformations map lines through the origin to lines through the origin, t maps the line joining the points $(0,0)$ and $(1,0)$ to the line joining the points $(0,0)$ and $(1,1)$, as illustrated in Figure 1; that is,

t maps the line $y = 0$ to the line $y = x$.

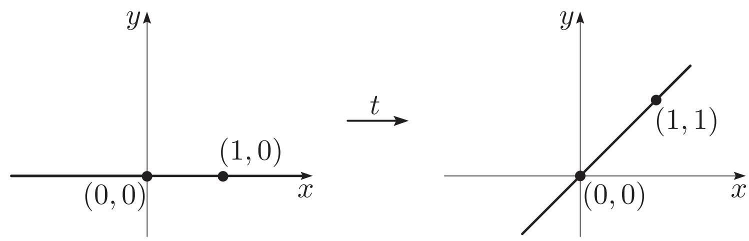


Figure 1 The image of the line $y = 0$ under the linear transformation t

Let us now calculate the image of the point $(1,-1)$:

$$t(1,-1) = (1 + 4(-1), 1 - 2(-1)) = (-3,3).$$

In this case, the linear transformation t maps the line joining the points $(0,0)$ and $(1,-1)$ to the line joining the points $(0,0)$ and $(-3,3)$, as illustrated in Figure 2; that is,

t maps the line $y = -x$ to itself.

Although t moves individual points on the line (except $(0,0)$) to other points, the line *as a whole* is unchanged.

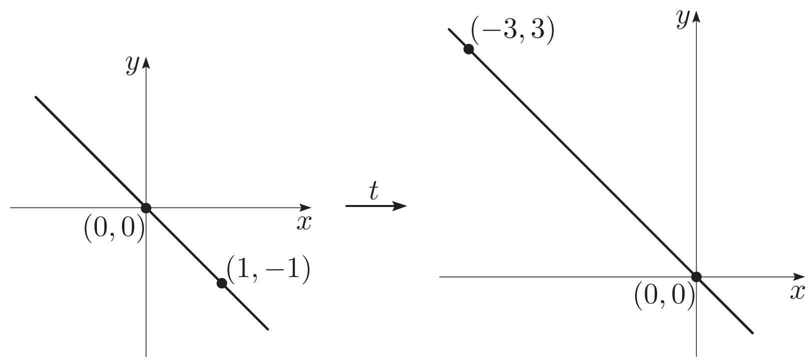


Figure 2 The image of the line $y = -x$ under the linear transformation t

The image of the point $(1,-1)$ under t is the point $(-3,3) = -3(1,-1)$. The vector $(1,-1)$ is scaled (stretched) by a factor of -3; that is, the resulting vector is three times the original magnitude and pointing in the opposite direction. In the next exercise you will investigate how other vectors lying along the line $y = -x$ are moved by t.

Exercise C115

For the above linear transformation t, calculate the images of the vectors $(2,-2)$ and $(-7,7)$. What do you notice?

We have seen that the linear transformation t scales some vectors lying along the line $y = -x$ by the factor -3. In fact this is true of any vector lying along this line, as we now show.

Let k be any real number, so that $(k, -k) = k(1, -1)$ is a vector lying along the line $y = -x$. Then

$$t(k, -k) = (k - 4k, k + 2k) = (-3k, 3k) = -3(k, -k),$$

which shows that t has the same scaling effect on each vector $(k, -k)$ lying along the line $y = -x$.

Does the linear transformation t map other lines through the origin to themselves?

Exercise C116

(a) For the above linear transformation t, calculate $t(0, 1)$, $t(1, 2)$ and $t(4, 1)$.

(b) Use one of the solutions to part (a) to write down another line in $\mathbb{R}^2$ that is mapped to itself by the linear transformation t.

(c) Find $t(4k, k)$.

We have seen that the linear transformation t maps each of the lines $y = -x$ and $x = 4y$ to itself. In both cases, each vector along the line is moved to a scalar multiple of itself: each vector lying along the line $y = -x$ is mapped to -3 times itself and each vector lying along the line $x = 4y$ is mapped to 2 times itself. We call the non-zero vectors lying along the line $y = -x$ *eigenvectors* of t with corresponding *eigenvalue* -3; for example, $(1, -1)$ and $(-7, 7)$ are eigenvectors of t with corresponding eigenvalue -3. Similarly, we call the non-zero vectors lying along the line $x = 4y$ eigenvectors of t with corresponding eigenvalue 2; for example, $(4, 1)$ and $(-8, -2)$ are eigenvectors of t with corresponding eigenvalue 2.

More generally, we make the following definitions; here and throughout this unit we use V to denote a finite-dimensional vector space.

Definitions

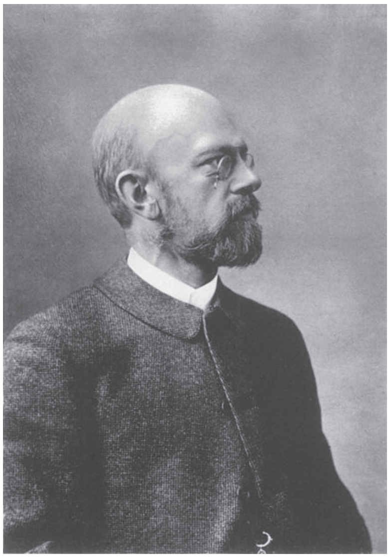
Let $t : V \longrightarrow V$ be a linear transformation. An **eigenvector** of t is a non-zero vector $\mathbf{v}$ that is mapped by t to a scalar multiple of itself; this scalar is the corresponding **eigenvalue**.

In symbols, a non-zero vector $\mathbf{v}$ is an eigenvector of a linear transformation t if

$$t(\mathbf{v}) = \lambda\mathbf{v}, \quad \text{for some } \lambda \in \mathbb{R};$$

λ is the corresponding eigenvalue.

We exclude the case $\mathbf{v} = \mathbf{0}$, since $t(\mathbf{0}) = \mathbf{0}$ for *every* linear transformation t. It is, however, possible for λ to be 0: when $\lambda = 0$, the linear transformation maps every vector corresponding to this eigenvalue to the origin – you will see an instance of this in Exercise C120.



David Hilbert

Eigen is a German word meaning *own*, *characteristic* or *special*. Another name for eigenvalue is *characteristic value*.

The *eigen* terms are associated with the German mathematician David Hilbert (1862–1943) who first used the terms *Eigenfunktion* (eigenfunction) and *Eigenwert* (Eigenvalue) in a series of papers on integral equations (1904–1910). It is possible that Hilbert was following the German physicist Hermann von Helmholtz (1821–1894) who used the term *Eigentöne* in acoustics in the nineteenth century.

In the 1920s the use of the eigen terminology was promoted through the development of the matrix mechanics formulation of quantum theory by the German physicist Werner Heisenberg (1901–1976) who wrote the new theory in the language of Hilbert and his followers.

Werner Heisenberg

In the example above we found two lines that are mapped to themselves by t, by considering the images of various points. This is a rather hit-and-miss way of finding eigenvalues and eigenvectors. Before developing a general method for finding them, we see that it is sometimes possible to do so by considering the geometry of the transformation.

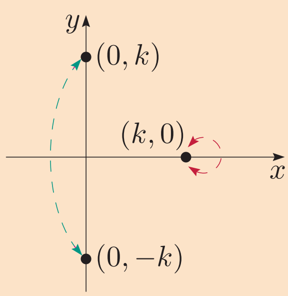
Worked Exercise C62

Let $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the linear transformation that maps each point to its reflection in the x-axis. By considering the geometric features of t, determine as many eigenvectors of t as you can and write down the corresponding eigenvalue in each case.

Solution

Reflection in the x-axis maps each point (x, y) to the point $(x, -y)$.

A sketch can help.



It therefore maps each point on the x-axis to itself, since

$$t(k,0) = (k,0) = 1(k,0),$$

so the vectors $(k,0)$, with $k \neq 0$, are eigenvectors of t with corresponding eigenvalue 1.

Note that $k \neq 0$, since we exclude the zero vector.

Similarly, each point on the y-axis is mapped to minus itself, since

$$t(0,k) = (0,-k) = -1(0,k),$$

so the vectors $(0,k)$, with $k \neq 0$, are eigenvectors of t with corresponding eigenvalue -1.

It is clear geometrically that every other line through the origin is changed by this transformation: we can find no other vector that maps to a multiple of itself. The only eigenvectors of t are therefore of the forms $(k,0)$ with eigenvalue 1 and $(0,k)$ with eigenvalue -1, where $k \neq 0$.

Exercise C117

By considering the geometric features of each of the following linear transformations of the plane, determine as many eigenvectors as you can and write down the corresponding eigenvalue in each case:

(a) reflection in the line $y = x$

(b) 2-dilation

(c) anticlockwise rotation through $\pi/2$ about the origin

(d) anticlockwise rotation through π about the origin.

In Exercise C117 it is possible to spot the eigenvectors geometrically. We now illustrate a general method to determine the eigenvalues and eigenvectors for any given transformation.

Consider again the linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$t(x,y) = (x + 4y, x - 2y).$$

We wish to find those vectors (x,y) that are mapped to scalar multiples of themselves; that is,

$$t(x,y) = \lambda(x,y) = (\lambda x, \lambda y).$$

We equate the expressions for $t(x,y)$ and obtain

$$(x + 4y, x - 2y) = (\lambda x, \lambda y).$$

Equating the first and second coordinates of these vectors, we obtain the system of linear equations

$$\begin{aligned} x + 4y &= \lambda x \\ x - 2y &= \lambda y. \end{aligned}$$

This is a system of two equations in the three unknowns x, y and λ. One way of solving this system is to move the terms on the right to the left-hand side. Thus we obtain the system

$$\begin{aligned} (1-\lambda)x + 4y &= 0 \\ x + (-2-\lambda)y &= 0. \end{aligned} \tag{1}$$

Equations (1) are called the *eigenvector equations*. We use them to find the possible values of λ, and then to find all the eigenvectors that correspond to these values. They are *homogeneous* equations in x and y since the constant terms are all zero.

Systems of homogeneous linear equations always have the trivial solution, in this case $x = 0$, $y = 0$, but this corresponds to the zero vector, which is excluded. Thus we seek *non-zero* solutions to the pair of homogeneous equations (1). Since we have two equations in three unknowns, such a system is bound to be dependant; that is, the homogeneous system has insufficient constraints on the unknowns to determine them uniquely.

From Theorem C19, Summary Theorem, in Unit C1 *Linear equations and matrices* we know that a homogeneous system has only the trivial solution if and only if the determinant of the coefficient matrix is non-zero. The contrapositive of this tells us that non-zero solutions exist if and only if the determinant of the coefficient matrix is 0; that is, if and only if

$$\begin{vmatrix} 1-\lambda & 4 \\ 1 & -2-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(1-\lambda)(-2-\lambda) - 4 = 0,$$

which simplifies (after some algebra) to

$$\lambda^2 + \lambda - 6 = 0.$$

This equation is called the *characteristic equation* of t, and its solutions are the eigenvalues we seek. Notice that the characteristic equation, whether or not it is written in terms of a determinant, is a polynomial equation in λ whose degree is the dimension of the domain of t – in this case 2. Here, we have

$$\lambda^2 + \lambda - 6 = (\lambda - 2)(\lambda + 3) = 0,$$

so the eigenvalues are $\lambda = 2$ and $\lambda = -3$.

To find the corresponding eigenvectors, we consider each eigenvalue λ in turn.

$\boxed{\lambda = 2}$ Putting $\lambda = 2$ into the eigenvector equations (1), we obtain

$$\begin{aligned} -x + 4y &= 0 \\ x - 4y &= 0. \end{aligned}$$

One equation is -1 times the other, so the equations are equivalent to the single equation

$$x = 4y.$$

Thus the eigenvectors corresponding to $\lambda = 2$ are the non-zero vectors (x, y) for which $x = 4y$; that is, the vectors of the form

$$(4k, k), \quad \text{where } k \neq 0.$$

Since we are working in a real vector space, in this case $\mathbb{R}^2$, when we are talking about eigenvectors, k represents a real number.

$\boxed{\lambda = -3}$ Putting $\lambda = -3$ into the eigenvector equations (1), we obtain

$$\begin{aligned} 4x + 4y &= 0 \\ x + y &= 0. \end{aligned}$$

These equations are equivalent to the single equation

$$y = -x.$$

Thus the eigenvectors corresponding to $\lambda = -3$ are the non-zero vectors (x, y) for which $y = -x$; that is, the vectors of the form

$$(k, -k), \quad \text{where } k \neq 0.$$

Thus the eigenvectors of t are the non-zero vectors of the following forms:

$(4k, k)$, corresponding to $\lambda = 2$,
$(k, -k)$, corresponding to $\lambda = -3$.

This method produces *all* the eigenvalues and eigenvectors of the linear transformation. On the other hand, trying to show that these are the only ones by calculating the images of various points, as we started to do at the beginning of the section, would take forever!

Exercise C118

Let $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the linear transformation given by

$$t(x, y) = (-5x + 3y, 6x - 2y).$$

(a) Find the eigenvector equations of t.

(b) Find the characteristic equation of t, and solve it to find the eigenvalues of t.

(c) Solve the eigenvector equations, for each eigenvalue in turn, to find the eigenvectors of t.

1.2 Finding eigenvalues and eigenvectors

You have just seen how to find the eigenvalues and eigenvectors of a given linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. This method, as it stands, is rather tedious to use to find eigenvalues and eigenvectors of linear transformations from $\mathbb{R}^3$ to $\mathbb{R}^3$, or $\mathbb{R}^4$ to $\mathbb{R}^4$, and so on. However, by introducing matrices, we can simplify the method.

We now work through the same example as in the previous subsection, but this time we use matrices.

Theorem C40 of Unit C3 tells us that there is a unique matrix for t with respect to the standard (ordered) basis in both the domain and codomain, and we use Strategy C15 from that unit to find this matrix. Recall that this strategy tells us essentially to 'read off' the matrix of a linear transformation when we are using the standard bases. We have $t(1,0) = (1,1)$ and $t(0,1) = (4,-2)$, so these vectors are the columns of the matrix of the linear transformation, since we are using the standard bases.

Therefore, with respect to the standard basis for $\mathbb{R}^2$, the linear transformation t given by $t(x,y) = (x+4y, x-2y)$ has the matrix representation

$$t : \mathbf{v} \longmapsto \mathbf{A}\mathbf{v}, \quad \text{where } \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} \text{ and } \mathbf{A} = \begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix}.$$

If $\mathbf{v}$ is an eigenvector of t with corresponding eigenvalue λ, then

$$t(\mathbf{v}) = \lambda\mathbf{v};$$

in matrix form, this becomes

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v};$$

that is,

$$\begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix}.$$

So

$$\begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} - \lambda\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2}$$

Using the 2×2 identity matrix $\mathbf{I}$, we can write

$$\lambda\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

so equation (2) can be written as

$$\left[\begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

that is,

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}.$$

We simplify this matrix equation and obtain

$$\begin{pmatrix} 1-\lambda & 4 \\ 1 & -2-\lambda \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This gives rise to the eigenvector equations

$$\begin{aligned} (1-\lambda)x + 4y &= 0 \\ x + (-2-\lambda)y &= 0, \end{aligned}$$

as before, which we labelled equations (1). The characteristic equation is

$$\begin{vmatrix} 1-\lambda & 4 \\ 1 & -2-\lambda \end{vmatrix} = 0,$$

that is,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

We can therefore find the characteristic equation directly from the matrix of the linear transformation (with respect to the standard basis for both the domain and codomain) by subtracting λ from each diagonal entry and then equating the determinant to zero.

Once we have found the eigenvalues, we use the same method as before to find the eigenvectors; that is, we substitute each eigenvalue in turn into the eigenvector equations and solve them.

In view of this connection with matrices, we adopt the following definitions.

Definitions

A non-zero vector $\mathbf{v}$ is an **eigenvector** of a square matrix $\mathbf{A}$ if

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \quad \text{for some } \lambda \in \mathbb{R};$$

λ is the corresponding **eigenvalue**.

The **characteristic equation** of a square matrix $\mathbf{A}$ is the equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

In this way we can refer to eigenvectors, eigenvalues and the characteristic equation of a matrix even when a linear transformation is not explicitly involved.

The matrix $\mathbf{A} - \lambda\mathbf{I}$ is obtained by subtracting λ from each entry on the diagonal of $\mathbf{A}$.

Eigenvalues and eigenvectors of matrices occur naturally in many applications – for example, in the study of vibrating mechanical systems. In such examples, the characteristic equation may have solutions that are not real numbers, and these *complex eigenvalues* have significance in these applications. In this unit we are primarily interested in linear transformations of the plane and of three-dimensional space, so complex eigenvalues play no role here: we are concerned only with *real* eigenvalues and eigenvectors.

Other areas of application include music, bridge design, oil exploration, image compression, and analysis of financial data. A particular example is the use of eigenvectors in the PageRank algorithm. This algorithm was invented by Larry Page and Sergey Brin, the founders of Google, in 1996 for use by the Google search engine to rank the importance of web pages. According to Google, PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites. The algorithm assigns a PageRank, or score, to each web page based on its linking web pages, with the links from different web pages being weighted according to particular criteria. The Google matrix represents the links between the web pages. A fundamental part of the algorithm is an iterative method that computes the dominant eigenvalue, that is, the eigenvalue of largest magnitude, and the corresponding eigenvector of the Google matrix to rank the web pages.

Larry Page and Sergey Brin

If a characteristic equation has no real solutions, then we say that there are no eigenvalues. For example, in Exercise C117(c), you considered the linear transformation representing an anticlockwise rotation through $\pi/2$ about the origin. The matrix of this linear transformation is

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

By the above definition, the characteristic equation of this linear transformation is

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$\lambda^2 + 1 = 0.$$

This equation has no real solutions: the linear transformation has no eigenvalues and hence no eigenvectors. This agrees with the geometric interpretation: no line through the origin is mapped to itself by this rotation.

We summarise this matrix method for finding eigenvalues and eigenvectors in the following strategy.

Strategy C18

To determine the eigenvalues and eigenvectors of a square matrix $\mathbf{A}$, do the following.

1. Find the eigenvalues:
   - write down the characteristic equation
     $$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$
   - expand this determinant to obtain a polynomial equation in λ
   - solve this equation to find the eigenvalues.
2. Find the eigenvectors:
   - write down the eigenvector equations
     $$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$
   - for each eigenvalue λ, solve this system of linear equations to find the corresponding eigenvectors.

We illustrate Strategy C18 with the following worked exercise and exercise.

Worked Exercise C63

Let $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the linear transformation given by

$$t(x, y) = (5x + 2y, 2x + 5y).$$

Write down the matrix of t with respect to the standard basis for $\mathbb{R}^2$, and find the eigenvalues and eigenvectors of t.

Solution

Since we are using the standard bases, we can simply 'read off' the matrix: the columns are the images of $(1, 0)$ and $(0, 1)$ under t.

The matrix of t with respect to the standard basis for $\mathbb{R}^2$ is

$$\mathbf{A} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}.$$

We use Strategy C18 to find the eigenvalues and eigenvectors of $\mathbf{A}$, which are the same as those of t.

First we find the eigenvalues of $\mathbf{A}$.

The characteristic equation of $\mathbf{A}$ is $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$; that is,

$$\begin{vmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(5 - \lambda)^2 - 4 = 0,$$

which simplifies to

$$\lambda^2 - 10\lambda + 21 = (\lambda - 7)(\lambda - 3) = 0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda = 7$ and $\lambda = 3$.

Next we find the eigenvectors of $\mathbf{A}$.

The eigenvector equations are $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$; that is,

$$\begin{pmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{0}$$

which we write as a system of linear equations.

The eigenvector equations are

$$\begin{aligned} (5-\lambda)x + \quad\;\; 2y &= 0 \\ 2x + (5-\lambda)y &= 0. \end{aligned}$$

$\boxed{\lambda = 7}$ The eigenvector equations become

$$\begin{aligned} -2x + 2y &= 0 \\ 2x - 2y &= 0. \end{aligned}$$

These equations are equivalent to the single equation

$$y = x.$$

Thus the eigenvectors corresponding to $\lambda = 7$ are the non-zero vectors for which $y = x$; that is, the vectors of the form

$$(k, k), \quad \text{where } k \neq 0.$$

$\boxed{\lambda = 3}$ The eigenvector equations become

$$\begin{aligned} 2x + 2y &= 0, \\ 2x + 2y &= 0. \end{aligned}$$

These equations are equivalent to the single equation

$$y = -x.$$

Thus the eigenvectors corresponding to $\lambda = 3$ are the non-zero vectors for which $y = -x$; that is, the vectors of the form

$$(k, -k), \quad \text{where } k \neq 0.$$

Thus the eigenvectors of the linear transformation t are the non-zero vectors of the following forms:

(k, k), corresponding to $\lambda = 7$,
$(k, -k)$, corresponding to $\lambda = 3$.

Exercise C119

For each of the following linear transformations $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, write down the matrix of t with respect to the standard basis for $\mathbb{R}^2$, and find the eigenvalues and eigenvectors of t.

(a) $t(x, y) = (x + 3y, 2x - 4y)$ (b) $t(x, y) = (x - 2y, -2x - 2y)$

So far we have concentrated on linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$ and on 2×2 matrices. We now use Strategy C18 to find the eigenvalues and eigenvectors of a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^3$ using a 3×3 matrix. Notice that here the characteristic equation is again a polynomial equation in λ whose degree is the dimension of the domain of t – in this case 3.

Worked Exercise C64

Let $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the linear transformation given by

$$t(x, y, z) = (2x + z, -x + 2y + 3z, x + 2z).$$

Write down the matrix of t with respect to the standard basis for $\mathbb{R}^3$, and find the eigenvalues and eigenvectors of t.

Solution

Since we are using the standard basis, we can again simply 'read off' the matrix: the columns are the images of $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ under t.

The matrix of t with respect to the standard basis for $\mathbb{R}^3$ is

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 0 & 2 \end{pmatrix}.$$

We use Strategy C18 to find the eigenvalues and eigenvectors of $\mathbf{A}$, which are the same as those of t.

First we find the eigenvalues of $\mathbf{A}$.

Here we need the 3×3 identity matrix $\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and so subtract λ from the three diagonal entries of $\mathbf{A}$.

The characteristic equation is $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$; that is,

$$\begin{vmatrix} 2-\lambda & 0 & 1 \\ -1 & 2-\lambda & 3 \\ 1 & 0 & 2-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(2-\lambda)\begin{vmatrix} 2-\lambda & 3 \\ 0 & 2-\lambda \end{vmatrix} - 0 + \begin{vmatrix} -1 & 2-\lambda \\ 1 & 0 \end{vmatrix} = 0.$$

Simplifying this expression, we obtain

$$(2-\lambda)((2-\lambda)^2 - 0) + (0 - (2-\lambda)) = 0.$$

When there is a common factor, it is best to keep this separate: the problem then reduces to factorising the remaining quadratic polynomial.

Taking out the common factor gives

$$(2-\lambda)((2-\lambda)^2 - 1) = 0,$$

which simplifies to

$$(2-\lambda)(\lambda^2 - 4\lambda + 3) = 0.$$

We can factorise this characteristic equation as

$$(2-\lambda)(\lambda-3)(\lambda-1) = 0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda = 3$, $\lambda = 2$ and $\lambda = 1$.

Next we find the eigenvectors of $\mathbf{A}$.

The eigenvector equations are

$$\begin{aligned} (2-\lambda)x \qquad\qquad &+ \qquad z = 0 \\ -x + (2-\lambda)y &+ \qquad 3z = 0 \\ x \qquad\qquad &+ (2-\lambda)z = 0. \end{aligned}$$

$\boxed{\lambda = 3}$ The eigenvector equations become

$$\begin{aligned} -x \qquad + \; z &= 0 \\ -x - y + 3z &= 0 \\ x \qquad - \; z &= 0. \end{aligned}$$

It may sometimes be necessary to use the method of Gauss–Jordan elimination from Unit C1, but here the solutions can be found directly.

The first and third equations imply that

$$z = x.$$

Substituting this into the second equation yields the equation

$$2x - y = 0.$$

Thus the eigenvectors corresponding to $\lambda = 3$ are the non-zero vectors (x, y, z) satisfying $z = x$ and $y = 2x$; that is, the vectors of the form

$$(k, 2k, k), \quad \text{where } k \neq 0.$$

$\boxed{\lambda = 2}$ The eigenvector equations become

$$\begin{aligned} z &= 0 \\ -x + 3z &= 0 \\ x \qquad &= 0. \end{aligned}$$

These equations have the solution

$$z = 0 \quad \text{and} \quad x = 0.$$

However, there are no constraints on the unknown y. Thus the eigenvectors corresponding to $\lambda = 2$ are the non-zero vectors (x, y, z) satisfying $x = 0$ and $z = 0$; that is, the vectors of the form

$$(0, k, 0), \quad \text{where } k \neq 0.$$

$\boxed{\lambda = 1}$ The eigenvector equations become

$$\begin{aligned} x \qquad + \quad z &= 0 \\ -x + y + 3z &= 0 \\ x \qquad + \quad z &= 0. \end{aligned}$$

The first and third equations imply that

$$z = -x.$$

Substituting this into the second equation yields the equation

$$-4x + y = 0.$$

Thus the eigenvectors corresponding to $\lambda = 1$ are the non-zero vectors (x, y, z) satisfying $z = -x$ and $y = 4x$; that is, the vectors of the form

$$(k, 4k, -k), \quad \text{where } k \neq 0.$$

Thus the eigenvectors of the linear transformation t are the non-zero vectors of the following forms:

$$\begin{aligned} &(k, 2k, k), \text{ corresponding to } \lambda = 3, \\ &(0, k, 0), \text{ corresponding to } \lambda = 2, \\ &(k, 4k, -k), \text{ corresponding to } \lambda = 1. \end{aligned}$$

Although cubic polynomials may not always be easy to factorise, you met some ways of factorising such polynomials in Subsection 1.4 of Unit A2 *Number systems*. However, we will usually deal with examples that factorise easily.

The following result, which we do not prove here, gives a useful check on the values found for the eigenvalues. You are asked to prove it yourself for 2×2 matrices in the additional exercises booklet for this unit.

Proposition C56

The sum of the eigenvalues of a square matrix $\mathbf{A}$ is equal to the sum of the diagonal entries of $\mathbf{A}$.

For example, in Worked Exercise C64 the eigenvalues are 3, 2 and 1, which sum to 6, and the diagonal entries of the matrix $\mathbf{A}$ are 2, 2 and 2, which also sum to 6.

The sum of the diagonal entries of a square matrix is sometimes referred to as the **trace** of the matrix.

Exercise C120

Let $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the linear transformation given by

$$t(x, y, z) = (4x + 2y, 2x + 3y + 2z, 2y + 2z).$$

Write down the matrix of t with respect to the standard basis for $\mathbb{R}^3$, and find the eigenvalues and eigenvectors of t.

In most of the examples we have seen so far, the eigenvalues have not been easy to recognise directly and Strategy C18 has been required to find them. This is not always the case, as the following exercise illustrates.

Exercise C121

Find the eigenvalues of each of the following matrices.

(a) $\begin{pmatrix} 1 & 2 \\ 0 & 6 \end{pmatrix}$ (b) $\begin{pmatrix} 8 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 21 \end{pmatrix}$ (c) $\begin{pmatrix} 4 & 0 & 0 \\ 25 & -2 & 0 \\ 17 & \pi & 6 \end{pmatrix}$

Finding eigenvalues of triangular and diagonal matrices is straightforward, as Exercise C121 illustrates. The eigenvalues are the diagonal entries of the matrix and no calculation is needed to find them.

Theorem C57

The eigenvalues of a triangular matrix and of a diagonal matrix are the diagonal entries of the matrix.

Proof A lower triangular matrix has every entry above the main diagonal zero. A diagonal matrix and the transpose of an upper triangular matrix are lower triangular matrices, so we can consider just lower triangular matrices here.

Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ lower triangular matrix, so $a_{ij} = 0$ for all $j > i$. The eigenvalues of $\mathbf{A}$ are the solutions to the characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. Now $\mathbf{A} - \lambda\mathbf{I}$ has diagonal entries $a_{ii} - \lambda$, and every entry above the main diagonal is zero.

We expand the determinant along the top row and continue by expanding along the top row of the resulting determinants until the only determinants in the expression are of size 2×2.

The first term in the full expansion of the determinant is the only non-zero term in the expansion because of the placement of the zeros in the smaller determinants. This non-zero term is $(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$. Therefore the solutions to the characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ are $a_{11}, a_{22}, \ldots, a_{nn}$, by the Factor Theorem (Theorem A2 in Unit A2), and the eigenvalues of $\mathbf{A}$ are precisely the diagonal entries of the matrix.

A diagonal matrix is lower triangular and $\det \mathbf{A}^T = \det \mathbf{A}$, so the eigenvalues of a triangular or diagonal matrix are the diagonal entries. ■

1.3 Eigenspaces

In Subsection 1.1 we considered the linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$t(x, y) = (x + 4y, x - 2y),$$

and saw that each of the lines $y = -x$ and $x = 4y$ is mapped to itself.

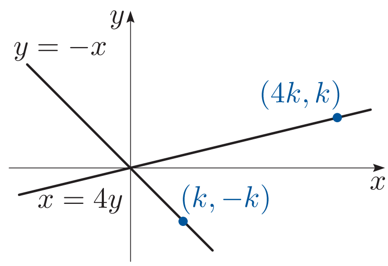


Figure 3 The lines comprising the eigenvectors of t

The line $y = -x$, shown in Figure 3, consists of the points of the form $(k, -k)$, each of which is an eigenvector of t corresponding to the eigenvalue $\lambda = -3$, except when $k = 0$, which is specifically excluded.

Similarly, the line $x = 4y$, also shown in Figure 3, consists of the points of the form $(4k, k)$, each of which is an eigenvector corresponding to the eigenvalue $\lambda = 2$, except when $k = 0$.

For each eigenvalue λ, if we look at *all* the solutions to the equation $t(\mathbf{v}) = \lambda\mathbf{v}$ (including $\mathbf{v} = \mathbf{0}$), then we obtain a line through the origin. The set of such solutions is a subspace of the domain of t.

Theorem C58

Let $t : V \longrightarrow V$ be a linear transformation. For each eigenvalue λ of t, let $S(\lambda)$ be the set of vectors satisfying $t(\mathbf{v}) = \lambda\mathbf{v}$; that is, $S(\lambda)$ is the set of eigenvectors corresponding to λ, together with the zero vector $\mathbf{0}$. Then $S(\lambda)$ is a subspace of V.

Proof Consider any eigenvalue λ of a linear transformation $t : V \longrightarrow V$.

We use Strategy C10 from Unit C2, *Vector spaces* and first check that $\mathbf{0} \in S(\lambda)$.

For any linear transformation t, we have $t(\mathbf{0}) = \mathbf{0} = \lambda\mathbf{0}$, so $\mathbf{0} \in S(\lambda)$.

Next we check that if $\mathbf{v}_1, \mathbf{v}_2 \in S(\lambda)$, then $\mathbf{v}_1 + \mathbf{v}_2 \in S(\lambda)$.

Let $\mathbf{v}_1, \mathbf{v}_2 \in S(\lambda)$. Then

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2) = \lambda\mathbf{v}_1 + \lambda\mathbf{v}_2 = \lambda(\mathbf{v}_1 + \mathbf{v}_2),$$

since t is a linear transformation.

Hence $\mathbf{v}_1 + \mathbf{v}_2 \in S(\lambda)$.

Finally, we check that if $\mathbf{v} \in S(\lambda)$ and $\alpha \in \mathbb{R}$, then $\alpha\mathbf{v} \in S(\lambda)$.

Let $\mathbf{v} \in S(\lambda)$ and $\alpha \in \mathbb{R}$. Then

$$t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v}) = \alpha\lambda\mathbf{v} = \lambda(\alpha\mathbf{v}),$$

since t is a linear transformation.

Hence $\alpha\mathbf{v} \in S(\lambda)$.

Thus $S(\lambda)$ is a subspace of V. ■

Since $S(\lambda)$ is a subspace comprising eigenvectors (and $\mathbf{0}$), we call it an *eigenspace*.

Definition

Let $t : V \longrightarrow V$ be a linear transformation and, for each eigenvalue λ of t, let $S(\lambda)$ be the set of vectors satisfying $t(\mathbf{v}) = \lambda\mathbf{v}$. Then $S(\lambda)$ is the **eigenspace** of t corresponding to the eigenvalue λ.

Worked Exercise C65

Let $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the linear transformation given by

$$t(x, y, z) = (4x + 2y, 2x + 3y + 2z, 2y + 2z).$$

Find the eigenspace $S(0)$ of t, specify a basis for it and state its dimension.

(You found the eigenvalues and eigenvectors of this linear transformation in Exercise C120.)

Solution

The non-zero vectors of the form $(k, -2k, 2k)$ are the eigenvectors of t corresponding to the eigenvalue $\lambda = 0$.

The eigenspace $S(0)$ is therefore the set of vectors

$$\{(k, -2k, 2k) : k \in \mathbb{R}\}.$$

Any vector in $S(0)$ can be written as $k(1,-2,2)$, so

$$\{(1,-2,2)\}$$

is a basis for $S(0)$. Thus $S(0)$ has dimension 1.

Geometrically, $S(0)$ is a line through the origin in the direction of the vector $(1,-2,2)$, so the only eigenvectors of t corresponding to $\lambda = 0$ are on this line.

Exercise C122

Let $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the linear transformation given by

$$t(x,y,z) = (4x+2y, 2x+3y+2z, 2y+2z).$$

Find the eigenspaces $S(6)$ and $S(3)$ of t. In each case, specify a basis and state the dimension of the eigenspace.

(In Exercise C120 you found that the eigenvectors of t are the non-zero vectors $(2k,2k,k)$ and $(-2k,k,2k)$, corresponding to the eigenvalues $\lambda = 6$ and $\lambda = 3$, respectively.)

Worked Exercise C66

Let $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the linear transformation given by

$$t(x,y,z) = (0,y,z).$$

Find all the eigenspaces of t. In each case, specify a basis and state the dimension of the eigenspace.

Solution

The matrix of t with respect to the standard basis for $\mathbb{R}^3$ is

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This matrix is diagonal, so the eigenvalues are the diagonal entries: $\lambda = 0$, $\lambda = 1$ and $\lambda = 1$.

The eigenvector equations are

$$\begin{aligned} -\lambda x \qquad\qquad\qquad\qquad &= 0 \\ (1-\lambda)y \qquad\qquad &= 0 \\ (1-\lambda)z &= 0. \end{aligned}$$

$\boxed{\lambda = 0}$ The eigenvector equations become

$$0x = 0,\ y = 0 \text{ and } z = 0.$$

Thus the eigenvectors corresponding to the eigenvalue $\lambda = 0$ are the non-zero vectors (x, y, z) satisfying $y = 0$ and $z = 0$; that is, the vectors of the form

$$(k, 0, 0), \quad \text{where } k \neq 0.$$

The eigenspace $S(0)$ is the set of vectors

$$\{(k, 0, 0) : k \in \mathbb{R}\}.$$

Any vector in $S(0)$ can be written as $k(1, 0, 0)$, so

$$\{(1, 0, 0)\}$$

is a basis for $S(0)$. Thus $S(0)$ has dimension 1.

Geometrically, $S(0)$ is the x-axis in $\mathbb{R}^3$.

$\boxed{\lambda = 1}$ The eigenvector equations reduce to the single equation

$$-x = 0.$$

Thus the eigenvectors corresponding to the eigenvalue $\lambda = 1$ are the non-zero vectors (x, y, z) satisfying $x = 0$; that is, the vectors of the form

$$(0, k, l), \quad \text{where } k \text{ and } l \text{ are not both } 0.$$

The eigenspace $S(1)$ is the set of vectors

$$\{(0, k, l) : k, l \in \mathbb{R}\}.$$

Any vector in $S(1)$ can be written as $k(0, 1, 0) + l(0, 0, 1)$, so

$$\{(0, 1, 0), (0, 0, 1)\}$$

is a basis for $S(1)$. Thus $S(1)$ has dimension 2.

Geometrically, $S(1)$ is the plane $x = 0$ through the origin.

In Worked Exercise C66 the (simplified) characteristic equation of the linear transformation t is

$$\lambda(\lambda - 1)^2 = 0.$$

The eigenvalue $\lambda = 1$ is a 'repeated' solution of this characteristic equation; it is a *multiple root* and we say that $\lambda = 1$ has *multiplicity* 2 because the factor $(\lambda - 1)$ occurs twice.

In general, we adopt the following definition.

Definition

If the characteristic equation of a square matrix $\mathbf{A}$ can be written as

$$(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_p)^{m_p} = 0,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_p$ are distinct, then the eigenvalue λ_j of $\mathbf{A}$ has **multiplicity** m_j, for $j = 1, 2, \ldots, p$.

For a triangular or diagonal matrix, the multiplicity of an eigenvalue is the number of times it appears on the main diagonal.

Exercise C123

Find the eigenvalues and eigenvectors of the matrix

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

For each eigenvalue λ, state its multiplicity, find the corresponding eigenspace $S(\lambda)$, specify a basis for $S(\lambda)$ and state its dimension.

From the examples that you have seen so far, you may be tempted to conjecture that the dimension of the eigenspace $S(\lambda)$, for a given eigenvalue λ, is equal to the multiplicity of λ. The following exercises give you the chance to investigate this conjecture.

Exercise C124

Find the eigenvalues and eigenvectors of the matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

For each eigenvalue λ, state its multiplicity, find the corresponding eigenspace $S(\lambda)$, specify a basis for $S(\lambda)$ and state its dimension.

Exercise C125

Find the eigenvalues and eigenvectors of the matrix

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & 4 & 1 \\ -1 & 1 & 4 \end{pmatrix}.$$

For each eigenvalue λ, state its multiplicity, find the corresponding eigenspace $S(\lambda)$, specify a basis for $S(\lambda)$ and state its dimension.

Hint: Look for factors in the characteristic equation and remember that $x^2 - 1 = (x - 1)(x + 1)$.

In Exercise C124 the eigenvalue $\lambda = 1$ has multiplicity 2, but it gives rise to an eigenspace of dimension only 1. In this case, the matrix represents a shear in the x-direction by a factor 1, as shown in Figure 4, and the only line through the origin left unchanged is the x-axis. Thus there is a single one-dimensional eigenspace, so the conjecture that the dimension of the eigenspace $S(\lambda)$ is equal to the multiplicity of λ is false.

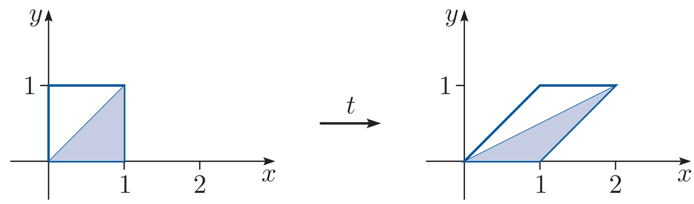


Figure 4 A shear in the x-direction by a factor 1

In Exercise C125 both eigenspaces have dimension 1 despite the eigenvalue 2 having multiplicity 2 and the eigenvalue 5 having multiplicity 1. In general, it can be shown that the dimension of an eigenspace cannot exceed the multiplicity of the corresponding eigenvalue, but we will not prove this.

2 Diagonalising matrices

In this section you will use the methods of finding eigenvalues and their corresponding eigenvectors that you met in the previous section to address the question posed in the introduction:

> Is it possible to find a basis for both the domain and codomain so that the matrix of a linear transformation is a diagonal matrix?

It is therefore important that you are confident with the material in Section 1 before starting to study this section.

2.1 Eigenvector bases

In Section 1 we introduced the notions of an eigenvalue λ and corresponding eigenvector $\mathbf{v}$ of a linear transformation $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$; that is, a non-zero vector $\mathbf{v}$ whose image $t(\mathbf{v})$ is $\lambda\mathbf{v}$. For example, in Exercise C119(a) you saw that the linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$t(x, y) = (x + 3y, 2x - 4y)$$

has eigenvalues $\lambda = -5$ and $\lambda = 2$ with corresponding eigenvectors the non-zero vectors of the forms $(k, -2k)$ and $(3k, k)$, respectively. We can choose any value of k ($k \neq 0$) to specify specific eigenvectors; here, putting $k = 1$ in both gives $(1, -2)$ and $(3, 1)$. Since $(3, 1)$ is not a multiple of $(1, -2)$, these two eigenvectors are linearly independent (this is the case whatever values of k are chosen). Therefore, by Theorem C25 in Unit C2, these linearly independent eigenvectors form a basis for $\mathbb{R}^2$ – the domain and codomain of t. We say that $\{(1, -2), (3, 1)\}$ is an *eigenvector basis* of t.

Definition

Let $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation and let E be a basis for $\mathbb{R}^n$ consisting of eigenvectors of t. The basis E is an **eigenvector basis** of t.

Exercise C126

Verify that $\{(-2,1),(1,2)\}$ is an eigenvector basis of the linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$t(x,y) = (x - 2y, -2x - 2y).$$

(In Exercise C119(b) you found that the eigenvectors of t are the non-zero vectors $(-2k,k)$ and $(k,2k)$, corresponding to the eigenvalues $\lambda = 2$ and $\lambda = -3$, respectively.)

Exercise C127

The set $E = \{(0,1,-1),(-2,1,0),(1,0,-1)\}$ is a basis for $\mathbb{R}^3$. Verify that E is an eigenvector basis of the linear transformation $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ given by

$$t(x,y,z) = (-x + 2y + 2z, 2x + 2y + 2z, -3x - 6y - 6z).$$

In Unit C3 you met Strategy C15 for finding the matrix representation of a linear transformation $t : V \longrightarrow W$ with respect to given bases E and F for the domain and codomain of t. In this subsection you will see that this matrix representation is particularly simple if $W = V$, E is an eigenvector basis of t and $F = E$.

Recall that if $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis for V, and $\mathbf{v}$ is a vector in V such that $\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n$, then the numbers $v_1, \ldots, v_n$ are the E-coordinates of $\mathbf{v}$, and $\mathbf{v}_E = (v_1, \ldots, v_n)_E$ is the E-coordinate representation of $\mathbf{v}$. If E is the standard basis for V, then we usually omit the suffix E.

We begin by rewriting Strategy C15 for the particular case when $W = V$ and $F = E$ (not necessarily an eigenvector basis).

Strategy C19 (Strategy C15 with $W = V$ and $F = E$)

To find the matrix $\mathbf{A}$ of a linear transformation $t : V \longrightarrow V$ with respect to the basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, do the following.

1. Find $t(\mathbf{e}_1), t(\mathbf{e}_2), \ldots, t(\mathbf{e}_n)$.
2. Find the E-coordinates of each of these image vectors.
3. Construct the matrix $\mathbf{A}$ column by column using the E-coordinates of $t(\mathbf{e}_j)$ to form column j, for $j = 1, 2, \ldots, n$.

In the next worked exercise we illustrate what happens when we find the matrix of a linear transformation t with respect to an eigenvector basis of t.

Worked Exercise C67

Consider the linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$t(x, y) = (x + 3y, 2x - 4y).$$

(a) Write down the matrix of t with respect to the standard basis for $\mathbb{R}^2$.

(b) Find the matrix of t with respect to the eigenvector basis

$$E = \{(1, -2), (3, 1)\}.$$

Solution

(a) The matrix of t with respect to the standard basis for $\mathbb{R}^2$ is

$$\begin{pmatrix} 1 & 3 \\ 2 & -4 \end{pmatrix}.$$

(b) Following Strategy C19, first we find the images of the vectors in the basis $E = \{(1, -2), (3, 1)\}$:

$$t(1, -2) = (-5, 10) \quad \text{and} \quad t(3, 1) = (6, 2).$$

We now write these image vectors in terms of their coordinates with respect to the eigenvector basis; that is, we express each of these vectors as a linear combination of the basis vectors $E = \{(1, -2), (3, 1)\}$. The resulting calculations are remarkably straightforward!

Next we find the E-coordinates of each of these image vectors:

$$\begin{aligned} (-5, 10) &= -5(1, -2) + 0(3, 1) \\ &= (-5, 0)_E, \\ (6, 2) &= 0(1, -2) + 2(3, 1) \\ &= (0, 2)_E. \end{aligned}$$

Therefore $t(1, -2) = (-5, 0)_E$ and $t(3, 1) = (0, 2)_E$. So the matrix of t with respect to the eigenvector basis E is

$$\begin{pmatrix} -5 & 0 \\ 0 & 2 \end{pmatrix}.$$

In Worked Exercise C67(b) we found that the matrix of t with respect to the eigenvector basis is diagonal and that its diagonal entries are the eigenvalues of the linear transformation t. This is because the matrix of the linear transformation t maps the basis vectors to their images under t, but these basis vectors are precisely the eigenvectors that get mapped to

multiples of themselves. You should find a similar outcome in the next exercise.

Exercise C128

Consider the linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$t(x, y) = (x - 2y, -2x - 2y).$$

(a) Write down the matrix of t with respect to the standard basis for $\mathbb{R}^2$.

(b) Find the matrix of t with respect to the eigenvector basis

$$E = \{(-2, 1), (1, 2)\},$$

which you found in Exercise C126.

Worked Exercise C67(b) and Exercise C128(b) are special cases of the following result. We use the letter $\mathbf{D}$ in this result because the matrix is diagonal.

Theorem C59

Let $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation, let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be an eigenvector basis of t and let $t(\mathbf{e}_j) = \lambda_j \mathbf{e}_j$, for $j = 1, 2, \ldots, n$. Then the matrix of t with respect to the eigenvector basis E is

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Proof Let t and E be as in the statement of the theorem. We use Strategy C19 to find the matrix of t with respect to the eigenvector basis E.

Eigenvector $\mathbf{e}_j$ corresponds to eigenvalue λ_j.

We have

$$t(\mathbf{e}_j) = \lambda_j \mathbf{e}_j, \quad \text{for } j = 1, 2, \ldots, n.$$

We find the E-coordinates of each of these image vectors:

$$\begin{aligned} t(\mathbf{e}_1) &= \lambda_1 \mathbf{e}_1 + 0\mathbf{e}_2 + \cdots + 0\mathbf{e}_n = (\lambda_1, 0, \ldots, 0)_E, \\ t(\mathbf{e}_2) &= 0\mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \cdots + 0\mathbf{e}_n = (0, \lambda_2, \ldots, 0)_E, \\ &\ \vdots \\ t(\mathbf{e}_n) &= 0\mathbf{e}_1 + 0\mathbf{e}_2 + \cdots + \lambda_n \mathbf{e}_n = (0, 0, \ldots, \lambda_n)_E. \end{aligned}$$

So the matrix of t with respect to the eigenvector basis E is

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

as claimed. ■

Using this result we can easily write down the matrix of a linear transformation with respect to an eigenvector basis.

Exercise C129

Consider the linear transformation $t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ given by

$$t(x, y, z) = (-x + 2y + 2z, 2x + 2y + 2z, -3x - 6y - 6z),$$

with eigenvector basis

$$E = \{(0, 1, -1), (-2, 1, 0), (1, 0, -1)\}.$$

Use the solution to Exercise C127 to write down the matrix of t with respect to this eigenvector basis.

2.2 Transition matrices

Suppose that $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear transformation and E is an eigenvector basis of t. We have just shown that the matrix of t with respect to the eigenvector basis E is a diagonal matrix $\mathbf{D}$.

Figures 5 and 6 show the linear transformation t with respect to the eigenvector basis E and the standard basis, respectively.

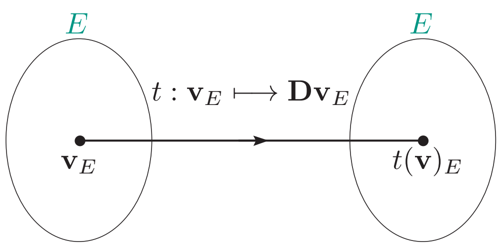


Figure 5 The linear transformation t with eigenvector basis E for the domain and codomain

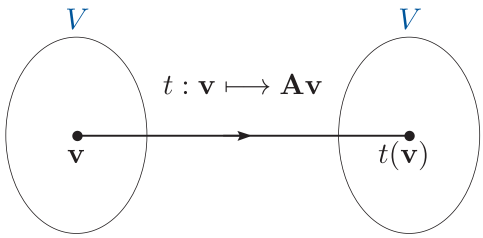


Figure 6 The linear transformation t with standard basis V for the domain and codomain

It is natural to ask whether there is any relationship between this matrix $\mathbf{D}$ and the matrix $\mathbf{A}$ of t with respect to the standard basis for $\mathbb{R}^n$. It turns out that there is an algebraic relationship between the matrices $\mathbf{D}$ and $\mathbf{A}$.

We now show this relationship. To do this, first we find an algebraic relationship between the E-coordinate representation of a vector $\mathbf{v}_E$ (as in Figure 5) and the standard coordinate representation of the same vector (as in Figure 6). We begin by doing this for the example that we considered at the beginning of the section, where $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is the linear transformation given by

$$t(x, y) = (x + 3y, 2x - 4y)$$

and E is the eigenvector basis $\{(1, -2), (3, 1)\}$.

Suppose that the E-coordinate representation of a vector $\mathbf{v}$ in $\mathbb{R}^2$ is

$$\mathbf{v}_E = (a, b)_E.$$

What are the standard coordinates of $\mathbf{v}$?

In column form,

$$\mathbf{v} = a\begin{pmatrix} 1 \\ -2 \end{pmatrix} + b\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} a + 3b \\ -2a + b \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}_E.$$

Thus in matrix form we have

$$\mathbf{v} = \mathbf{P}\mathbf{v}_E,$$

where

$$\mathbf{P} = \begin{pmatrix} 1 & 3 \\ -2 & 1 \end{pmatrix}.$$

Now, by the Summary Theorem (Theorem C19 in Unit C1), a square matrix is invertible if and only if its determinant is non-zero. Here we have $\det \mathbf{P} = 1 - (-6) = 7 \neq 0$, so $\mathbf{P}$ is invertible with inverse $\mathbf{P}^{-1}$.

Since $\mathbf{v} = \mathbf{P}\mathbf{v}_E$, it follows that

$$\mathbf{P}^{-1}\mathbf{v} = \mathbf{P}^{-1}(\mathbf{P}\mathbf{v}_E) = (\mathbf{P}^{-1}\mathbf{P})\mathbf{v}_E = \mathbf{v}_E.$$

So multiplication on the left by the matrix $\mathbf{P}$ converts the E-coordinate representation of a vector into the standard coordinate representation and, similarly, multiplication on the left by the matrix $\mathbf{P}^{-1}$ converts the standard coordinate representation of a vector into the E-coordinate representation.

In this case the columns of $\mathbf{P}$ are formed from the standard coordinates of the vectors in E, but this is no coincidence. This simple relationship between the matrix $\mathbf{P}$ and the basis E always holds and we call $\mathbf{P}$ the *transition matrix* from the basis E to the standard basis for $\mathbb{R}^2$.

The general definition is as follows.

Definition

Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be a basis for $\mathbb{R}^n$. The **transition matrix P** from the basis E to the standard basis for $\mathbb{R}^n$ is the matrix whose jth column is formed from the standard coordinates of $\mathbf{e}_j$.

Exercise C130

(a) Write down the transition matrix $\mathbf{P}$ from the basis $E = \{(1,3),(2,5)\}$ to the standard basis for $\mathbb{R}^2$.

(b) Write down the transition matrix $\mathbf{P}$ from the basis $E = \{(0,1,-1), (-2,1,0),(1,0,-1)\}$ to the standard basis for $\mathbb{R}^3$.

In the example above, we have seen that the transition matrix $\mathbf{P}$ from the basis $E = \{(1,-2),(3,1)\}$ to the standard basis for $\mathbb{R}^2$ converts E-coordinate representations into standard coordinate representations, and that $\mathbf{P}^{-1}$ converts standard coordinate representations into E-coordinate representations. This is true in general.

Theorem C60

Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be a basis for $\mathbb{R}^n$ and let $\mathbf{P}$ be the transition matrix from the basis E to the standard basis for $\mathbb{R}^n$. Then the standard coordinate representation of a vector in $\mathbb{R}^n$ is given by

$$\mathbf{v} = \mathbf{P}\mathbf{v}_E.$$

Moreover, $\mathbf{P}$ is invertible and

$$\mathbf{v}_E = \mathbf{P}^{-1}\mathbf{v}.$$

Proof The matrix $\mathbf{P}$ converts the E-coordinate representation of a vector in $\mathbb{R}^n$ to the standard coordinate representation of the *same* vector in $\mathbb{R}^n$, so in effect it is the matrix of the identity linear transformation $i : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with respect to the basis E in the domain and the standard basis in the codomain.

The statement $\mathbf{v} = \mathbf{P}\mathbf{v}_E$ is equivalent to the statement that $\mathbf{P}$ is the matrix of the identity transformation i of $\mathbb{R}^n$ with respect to the basis E for the domain and the standard basis for the codomain.

To find this matrix $\mathbf{P}$, we use Strategy C15 from Unit C3. We begin by finding the images under i of the vectors in the domain basis E:

$$i(\mathbf{e}_1) = \mathbf{e}_1, \quad i(\mathbf{e}_2) = \mathbf{e}_2, \quad \ldots, \quad i(\mathbf{e}_n) = \mathbf{e}_n.$$

It now follows from Strategy C15 that each column of $\mathbf{P}$ is formed from the standard coordinates of the corresponding basis vector, so $\mathbf{P}$ is the transition matrix from the basis E to the standard basis for $\mathbb{R}^n$, as claimed.

We know that the identity transformation i is invertible and that $i^{-1} = i$. It follows from the Inverse Rule (Theorem C45 in Unit C3) that $\mathbf{P}$ is invertible and that $\mathbf{P}^{-1}$ is the matrix of $i : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with respect to the standard basis for the domain and the basis E for the codomain; that is,

$$\mathbf{v} \longmapsto \mathbf{v}_E = \mathbf{P}^{-1}\mathbf{v}.$$ ■

When E is the standard basis for $\mathbb{R}^n$, the matrix $\mathbf{P}$ is the identity matrix $\mathbf{I}_n$, as you would expect.

We also get the following corollary from Theorem C60.

Corollary C61

The rows or columns of an $n \times n$ matrix $\mathbf{A}$ form a set of n linearly independent vectors if and only if $\det \mathbf{A} \neq 0$.

Proof Let $\mathbf{A}$ be an $n \times n$ matrix.

We start by proving the *only if* part.

We first show that if the columns of $\mathbf{A}$ are linearly independent, then $\det \mathbf{A} \neq 0$.

Suppose the columns are linearly independent, then the columns form a basis for $\mathbb{R}^n$ and $\mathbf{A}$ is the transition matrix from this basis to the standard basis. Hence $\mathbf{A}$ is invertible by Theorem C60, and so $\det \mathbf{A} \neq 0$ by the Summary Theorem (Theorem C19 in Unit C1).

If the rows of $\mathbf{A}$ are linearly independent then we consider the transpose $\mathbf{A}^T$.

Suppose the rows of $\mathbf{A}$ are linearly independent, then the columns of $\mathbf{A}^T$ are linearly independent and $\det \mathbf{A}^T \neq 0$ by the above reasoning. We have $\det \mathbf{A} = \det \mathbf{A}^T$ by Theorem C14 in Unit C1, and hence $\det \mathbf{A} \neq 0$, as required.

We now prove the *if* part using the contrapositive; that is, we show that if the rows or columns of $\mathbf{A}$ are *not* linearly independent then $\det \mathbf{A} = 0$.

Suppose the rows of $\mathbf{A}$ form a linearly dependent set, then the row-reduced form of $\mathbf{A}$ contains a zero row, so $\mathbf{A}$ is not invertible by the Invertibility Theorem (Theorem C7 in Unit C1), and hence $\det \mathbf{A} = 0$ by the Summary Theorem.

Suppose the columns of $\mathbf{A}$ form a linearly dependent set, then the rows of $\mathbf{A}^T$ are linearly dependent and $\det \mathbf{A} = \det \mathbf{A}^T = 0$ by the above reasoning.

Hence, if $\det \mathbf{A} \neq 0$, then the rows or columns of $\mathbf{A}$ form a linearly independent set of vectors. ■

Recall that our aim in this subsection is to relate the matrices $\mathbf{D}$ and $\mathbf{A}$, where $\mathbf{D}$ is the matrix of a linear transformation $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with respect to an eigenvector basis of t, and $\mathbf{A}$ is the matrix of t with respect to the standard basis for $\mathbb{R}^n$. Figure 7 shows how we can do this by using the transition matrix $\mathbf{P}$ from the eigenvector basis E to the standard basis for $\mathbb{R}^n$, so linking together Figures 5 and 6.

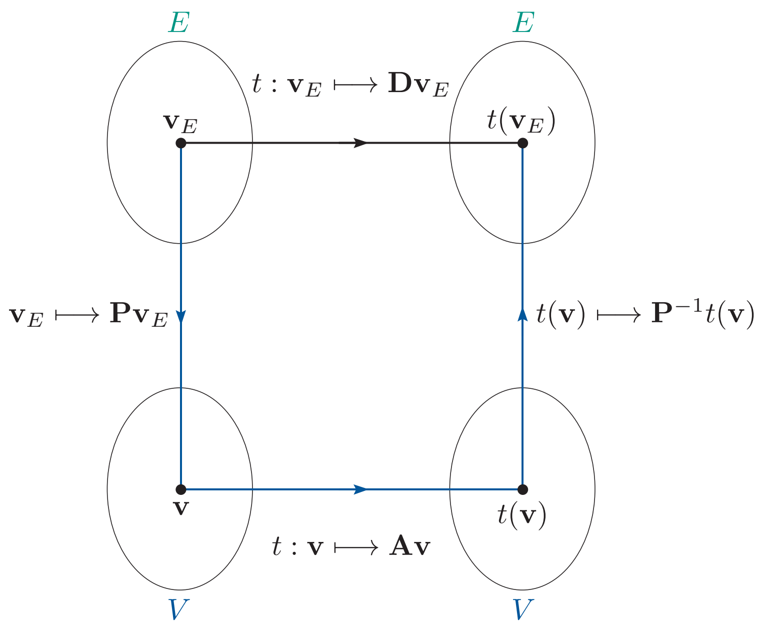


Figure 7 The transition matrix $\mathbf{P}$ from the eigenvector basis E of t to the standard basis for $\mathbb{R}^n$

The top line of the diagram shows that multiplication by $\mathbf{D}$ converts the E-coordinate representation of $\mathbf{v}$ to the E-coordinate representation of $t(\mathbf{v})$:

$$t(\mathbf{v})_E = \mathbf{D}\mathbf{v}_E. \tag{3}$$

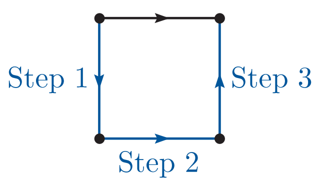


Figure 8 The transition in three steps

The diagram also shows that this change can be achieved in another way, in three steps, highlighted in Figure 8.

1. Use the transition matrix $\mathbf{P}$ to convert the E-coordinate representation of $\mathbf{v}$ to the standard coordinate representation of $\mathbf{v}$:

   $$\mathbf{v} = \mathbf{P}\mathbf{v}_E.$$

2. Multiply $\mathbf{v}$ on the left by matrix $\mathbf{A}$ to obtain the standard coordinate representation of $t(\mathbf{v})$:

   $$t(\mathbf{v}) = \mathbf{A}\mathbf{v} = \mathbf{A}\mathbf{P}\mathbf{v}_E.$$

3. Use the matrix $\mathbf{P}^{-1}$ to convert the standard coordinate representation of $t(\mathbf{v})$ to the E-coordinate representation of $t(\mathbf{v})$:

$$t(\mathbf{v})_E = \mathbf{P}^{-1}t(\mathbf{v}) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{v}_E.$$

Comparing this last equation with equation (3), we see that $\mathbf{D}$, $\mathbf{A}$ and $\mathbf{P}$ are related by the equation

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

Thus we have proved the following result.

Theorem C62

Let $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation and let E be an eigenvector basis of t. Let $\mathbf{A}$ be the matrix of t with respect to the standard basis for $\mathbb{R}^n$, let $\mathbf{D}$ be the matrix of t with respect to the eigenvector basis E and let $\mathbf{P}$ be the transition matrix from E to the standard basis for $\mathbb{R}^n$. Then

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

In fact, Theorem C62 holds for *any* basis E for $\mathbb{R}^n$, although $\mathbf{D}$ is diagonal only when E is an eigenvector basis.

Since $\mathbf{D}$, $\mathbf{A}$, $\mathbf{P}$ and $\mathbf{P}^{-1}$ are all square $n \times n$ matrices, we can multiply $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ on the left by the matrix $\mathbf{P}$ and on the right by the matrix $\mathbf{P}^{-1}$ to obtain the related equation

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}.$$

This algebraic relationship $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ may remind you of the algebraic relationship

$$y = g \circ x \circ g^{-1}$$

between *conjugate* permutations x and y in the symmetric group S_n, which you met in Subsection 4.1 of Unit B3. You saw in Unit C1 that the set of square invertible $n \times n$ matrices form a group under multiplication, and here the change of basis is in some sense equivalent to the 'renaming' in permutations. The matrices $\mathbf{D}$ and $\mathbf{A}$ are *conjugate matrices*: we will not use this concept here, but you will meet this idea of conjugacy in groups again in Book E.

We end this subsection by applying Theorem C62 to some examples.

Consider the linear transformation $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by

$$t(x, y) = (x + 3y, 2x - 4y).$$

In Worked Exercise C67 you saw that

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & -4 \end{pmatrix}$$

is the matrix of t with respect to the standard basis for $\mathbb{R}^2$ and that

$$\mathbf{D} = \begin{pmatrix} -5 & 0 \\ 0 & 2 \end{pmatrix}$$

is the matrix of t with respect to the eigenvector basis $E = \{(1, -2), (3, 1)\}$.

At the beginning of this subsection you saw that the transition matrix from the basis E to the standard basis for $\mathbb{R}^2$ is

$$\mathbf{P} = \begin{pmatrix} 1 & 3 \\ -2 & 1 \end{pmatrix}.$$

Now, using Strategy C4 from Unit C1, we have

$$\mathbf{P}^{-1} = \frac{1}{7}\begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{7} & -\frac{3}{7} \\ \frac{2}{7} & \frac{1}{7} \end{pmatrix},$$

so

$$\begin{aligned} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} &= \begin{pmatrix} \frac{1}{7} & -\frac{3}{7} \\ \frac{2}{7} & \frac{1}{7} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -5 & 0 \\ 0 & 2 \end{pmatrix} \\ &= \mathbf{D}, \end{aligned}$$

as claimed.

Exercise C131

Use the solution to Exercise C128 to find a matrix $\mathbf{P}$ such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}.$$

2.3 Diagonalisation

In this subsection you will consider the problem of determining when a matrix is *diagonalisable* and how to *diagonalise* a matrix when it is possible.

Definition

The matrix $\mathbf{A}$ is **diagonalisable** if there exists an invertible matrix $\mathbf{P}$ such that the matrix

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

is diagonal.

Clearly the matrices $\mathbf{A}$, $\mathbf{D}$ and $\mathbf{P}$ must all be square matrices of the same size.

If a matrix $\mathbf{A}$ is diagonalisable, then to diagonalise it we need to find both the diagonal matrix $\mathbf{D}$ *and* the invertible matrix $\mathbf{P}$, since it is this transition matrix $\mathbf{P}$ that links the matrix $\mathbf{A}$ with the diagonal matrix $\mathbf{D}$.

One particular use of diagonalisation of matrices is to find powers of matrices. We saw earlier that multiplying $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$ on the left by $\mathbf{P}$ and on the right by $\mathbf{P}^{-1}$ gives $\mathbf{A} = \mathbf{PDP}^{-1}$. Now consider powers of $\mathbf{A}$,

$$\begin{aligned}\mathbf{A}^2 &= (\mathbf{PDP}^{-1})(\mathbf{PDP}^{-1})\\ &= \mathbf{PD}(\mathbf{P}^{-1}\mathbf{P})\mathbf{DP}^{-1}\\ &= \mathbf{PDDP}^{-1}\\ &= \mathbf{PD}^2\mathbf{P}^{-1},\end{aligned}$$

and, in general we have

$$\mathbf{A}^n = \mathbf{PD}^n\mathbf{P}^{-1}, \quad \text{for } n = 1, 2, \ldots.$$

This last equation is useful for calculating powers of matrices, since calculating the nth power of a diagonal matrix is particularly simple: you need to find only the nth power of each diagonal entry. But first we need to be able to find both $\mathbf{D}$ and $\mathbf{P}$ (from which we can find $\mathbf{P}^{-1}$).

Exercise C132

(a) Write down $\mathbf{D}^5$, where $\mathbf{D} = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$.

(b) Calculate $\mathbf{A}^5$, where $\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}$.

(In Exercise C131 you found that $\mathbf{P} = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$ satisfies $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$.)

If $\mathbf{A}$ is any $n \times n$ matrix, then we can define a linear transformation t as:

$$\begin{aligned} t : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \mathbf{v} &\longmapsto \mathbf{Av}. \end{aligned}$$

In Section 1 we said that $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ with corresponding eigenvalue λ if $\mathbf{Av} = t(\mathbf{v}) = \lambda\mathbf{v}$; that is, if $\mathbf{v}$ is an eigenvector of t.

Definition

Let $\mathbf{A}$ be an $n \times n$ matrix and let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be a basis for $\mathbb{R}^n$ consisting of eigenvectors of $\mathbf{A}$. The basis E is an **eigenvector basis** of $\mathbf{A}$.

Thus E is an eigenvector basis of $\mathbf{A}$ if E is an eigenvector basis of t.

Worked Exercise C68

Find an eigenvector basis of the matrix

$$\mathbf{A} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}.$$

Solution

We found the eigenvectors of $\mathbf{A}$ in Worked Exercise C63.

The eigenvectors of $\mathbf{A}$ are the non-zero vectors of the following forms:

(k, k), corresponding to the eigenvalue $\lambda = 7$,

$(k, -k)$, corresponding to the eigenvalue $\lambda = 3$.

Since $(1, 1)$ and $(1, -1)$ are eigenvectors of $\mathbf{A}$, and $(1, -1)$ is not a multiple of $(1, 1)$, the set $E = \{(1, 1), (1, -1)\}$ is an eigenvector basis of $\mathbf{A}$.

Suppose that E is an eigenvector basis of the $n \times n$ matrix $\mathbf{A}$; that is, E is an eigenvector basis of the linear transformation $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ given by

$$t(\mathbf{v}) = \mathbf{A}\mathbf{v}.$$

It follows from Theorems C59 and C62 that if $\mathbf{P}$ is the transition matrix from the basis E to the standard basis for $\mathbb{R}^n$, then

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

is diagonal; that is, $\mathbf{A}$ is diagonalisable. This gives the following strategy for diagonalising a matrix, when this is possible.

Strategy C20

To diagonalise an $n \times n$ matrix $\mathbf{A}$:

1. find all the eigenvalues of $\mathbf{A}$
2. find (if possible) an eigenvector basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathbf{A}$
3. write down the transition matrix $\mathbf{P}$ whose jth column is formed from the standard coordinates of $\mathbf{e}_j$.

Then

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_j is the eigenvalue corresponding to the eigenvector $\mathbf{e}_j$.

The order of the eigenvalues down the diagonal of $\mathbf{D}$ must match the order of the eigenvectors in the basis E used to construct the transition matrix $\mathbf{P}$. When asked to diagonalise a matrix, it is not enough to write down a diagonal matrix containing the eigenvalues: you must also give the transition matrix $\mathbf{P}$.

The complexity involved in finding an eigenvector basis of $\mathbf{A}$ in step 2 of Strategy C20 depends on the matrix $\mathbf{A}$. In Worked Exercise C68 we formed an eigenvector basis of $\mathbf{A}$ by taking one eigenvector corresponding to each eigenvalue, ensuring that the eigenvectors were linearly independent. In general, we have the following result, which we will prove at the end of this subsection after looking at how it can be used. This result means that any eigenvector can be chosen for each (distinct) eigenvalue and there is no need to check that they are linearly independent.

Theorem C63

Let $\mathbf{A}$ be an $n \times n$ matrix with *distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and corresponding eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Then $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an eigenvector basis of $\mathbf{A}$.

We give an example of how Theorem C63 can be used.

Worked Exercise C69

Diagonalise the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 0 & 2 \end{pmatrix}.$$

Solution

We use Strategy C20.

We found the eigenvalues and eigenvectors of $\mathbf{A}$ in Worked Exercise C64.

The eigenvalues of $\mathbf{A}$ are $\lambda = 3$, $\lambda = 2$ and $\lambda = 1$.

The eigenvectors of $\mathbf{A}$ are the non-zero vectors of the following forms:

$$\begin{aligned} &(k, 2k, k), \text{ corresponding to } \lambda = 3, \\ &(0, k, 0), \text{ corresponding to } \lambda = 2, \\ &(k, 4k, -k), \text{ corresponding to } \lambda = 1. \end{aligned}$$

It follows from Theorem C63 that we can form an eigenvector basis of $\mathbf{A}$ by taking one eigenvector corresponding to each of the three distinct eigenvalues. For example,

$$E = \{(1,2,1),(0,1,0),(1,4,-1)\}$$

is an eigenvector basis of $\mathbf{A}$.

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 4 \\ 1 & 0 & -1 \end{pmatrix}.$$

Remember that the eigenvalues in $\mathbf{D}$ must appear in the same order as the corresponding eigenvectors in $\mathbf{P}$.

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If the eigenvectors had been chosen in a different order, then the order of the columns of the transition matrix $\mathbf{P}$ and the order of the diagonal entries of the resulting matrix $\mathbf{D}$ would have been different.

In addition, other transition matrices arise from using different eigenvectors for the eigenvector basis.

Another solution is

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad \text{where} \quad \mathbf{P} = \begin{pmatrix} 0 & -1 & 1 \\ 2 & -4 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

Both the order of the eigenvalues, and the eigenvectors chosen for the columns of P, differ here.

Exercise C133

Diagonalise the matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

(In Exercise C120 you found that the eigenvectors of $\mathbf{A}$ are the non-zero vectors $(2k,2k,k)$, $(-2k,k,2k)$ and $(k,-2k,2k)$, corresponding to the eigenvalues $\lambda = 6$, $\lambda = 3$ and $\lambda = 0$, respectively.)

It may be possible to find an eigenvector basis of an $n \times n$ matrix $\mathbf{A}$ even when $\mathbf{A}$ does not have n distinct eigenvalues.

Strategy C21

To find an eigenvector basis of an $n \times n$ matrix $\mathbf{A}$:

1. find a basis for each eigenspace of $\mathbf{A}$
2. form the set E of all the basis vectors found in step 1.

If there are n vectors in E, then E is an eigenvector basis of $\mathbf{A}$; otherwise E is not a basis.

The fact that E, as found in Strategy C21, is an eigenvector basis of $\mathbf{A}$ if and only if there are n vectors in E, can be proved in a similar way to Theorem C63, but the details are more complicated.

Worked Exercise C70

Diagonalise the matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

given that the eigenvalues of $\mathbf{A}$ are $\lambda = 8$, $\lambda = 2$ and $\lambda = 2$.

Solution

There are many possible solutions to this, and to each of the remaining exercises in this section.

We use Strategies C20 and C21, but start with the second part of Strategy C18. In general we will not list all the strategies involved.

To find the eigenspaces of $\mathbf{A}$, we consider the eigenvector equations

$$\begin{aligned} (4-\lambda)x + \quad 2y + \quad 2z &= 0 \\ 2x + (4-\lambda)y + \quad 2z &= 0 \\ 2x + \quad 2y + (4-\lambda)z &= 0, \end{aligned}$$

for each eigenvalue.

$\boxed{\lambda = 8}$ The eigenvector equations become

$$\begin{aligned} -4x + 2y + 2z &= 0 \\ 2x - 4y + 2z &= 0 \\ 2x + 2y - 4z &= 0. \end{aligned}$$

Subtracting the second equation from the first, we obtain $-6x + 6y = 0$, which implies that $x = y$. Substituting this into the third equation, we obtain $4x - 4z = 0$, which implies that $x = z$.

Thus $S(8) = \{(k, k, k) : k \in \mathbb{R}\}$.

$\boxed{\lambda = 2}$ All three eigenvector equations become

$$2x + 2y + 2z = 0,$$

that is, $x + y + z = 0$, so $z = -(x + y)$.

Thus $S(2) = \{(k, l, -(k + l)) : k, l \in \mathbb{R}\}$.

Any vector in $S(8)$ can be written as $k(1, 1, 1)$, and any vector in $S(2)$ can be written as $k(1, 0, -1) + l(0, 1, -1)$.

A basis for $S(8)$ is $\{(1, 1, 1)\}$ and a basis for $S(2)$ is $\{(1, 0, -1), (0, 1, -1)\}$. The set

$$E = \{(1, 1, 1), (1, 0, -1), (0, 1, -1)\}$$

contains three vectors, so it is an eigenvector basis of $\mathbf{A}$.

Note that Strategy C21 does not require us to prove linear independence of the vectors in E: combining the bases of the eigenspaces $S(2)$ and $S(8)$ gives a set of linearly independent vectors.

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}.$$

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Exercise C134

Diagonalise the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

If the matrix $\mathbf{A}$ does *not* have an eigenvector basis, then these methods cannot be applied and the matrix $\mathbf{A}$ is not diagonalisable – there is no transition matrix. For example, in Exercise C124 you saw that all the eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

are non-zero vectors of the form $(k, 0)$. Any two eigenvectors of $\mathbf{A}$ are linearly dependent, so there is no eigenvector basis. Thus there is no transition matrix and $\mathbf{A}$ is not diagonalisable.

Similarly the matrix

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 4 & 1 \\ -1 & 1 & 4 \end{pmatrix}$$

from Exercise C125 is also not diagonalisable. The eigenvectors corresponding to the eigenvalue $\lambda = 2$ of multiplicity 2 are the non-zero vectors of the form $(k, -k, k)$, so any two eigenvectors of $\mathbf{B}$ in $S(2)$ are linearly dependent. The other eigenvalue $\lambda = 5$ has multiplicity 1. As stated at the end of Section 1, the dimension of an eigenspace cannot exceed the multiplicity of the corresponding eigenvalue, and so there cannot be two linearly independent eigenvectors corresponding to eigenvalue $\lambda = 5$.

Therefore there is no set of three linearly independent eigenvectors and thus no eigenvector basis; there is no transition matrix and thus $\mathbf{B}$ is not diagonalisable.

We have shown that, if the matrix $\mathbf{A}$ of a linear transformation t has an eigenvector basis, then using this basis for both the domain and codomain results in a matrix of t that is a diagonal matrix. On the other hand, if there is an eigenvalue of multiplicity m for which there are fewer than m linearly independent eigenvectors, then there is no eigenvector basis and matrix $\mathbf{A}$ is not diagonalisable.

We end this section by proving Theorem C63 as promised.

Theorem C63

Let $\mathbf{A}$ be an $n \times n$ matrix with *distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and corresponding eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Then $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an eigenvector basis of $\mathbf{A}$.

Proof Let $\mathbf{A}$ and E be as in the statement of the theorem.

Since any linearly independent set of n vectors in $\mathbb{R}^n$ is a basis for $\mathbb{R}^n$, by Theorem C25 in Unit C2, we need show only that E is linearly independent. To do this, we assume that E is linearly dependent and obtain a contradiction.

If E is linearly independent, then E must be an eigenvector basis of $\mathbf{A}$.

Suppose to the contrary that E is linearly dependent. Then we can take the smallest value of m ($2 \leq m \leq n$) for which a set of m vectors in E is linearly dependent. By relabelling the eigenvectors (if necessary), we can write

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_m\mathbf{e}_m = \mathbf{0}, \tag{4}$$

with $\alpha_1 \neq 0$, $\alpha_2 \neq 0$, $\ldots$, $\alpha_m \neq 0$.

Multiplying both sides of equation (4) by matrix $\mathbf{A}$, we obtain

$$\mathbf{A}(\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_m\mathbf{e}_m) = \mathbf{A0},$$

that is,

$$\alpha_1\mathbf{Ae}_1 + \alpha_2\mathbf{Ae}_2 + \cdots + \alpha_m\mathbf{Ae}_m = \mathbf{0}.$$

Now, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$ are eigenvectors of $\mathbf{A}$ with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, so $\mathbf{Ae}_j = \lambda_j\mathbf{e}_j$ and

$$\alpha_1\lambda_1\mathbf{e}_1 + \alpha_2\lambda_2\mathbf{e}_2 + \cdots + \alpha_m\lambda_m\mathbf{e}_m = \mathbf{0}. \tag{5}$$

We now eliminate the vector $\mathbf{e}_m$. To do this, we multiply equation (4) by λ_m and subtract the result from equation (5):

$$\alpha_1(\lambda_1 - \lambda_m)\mathbf{e}_1 + \alpha_2(\lambda_2 - \lambda_m)\mathbf{e}_2 + \cdots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)\mathbf{e}_{m-1} = \mathbf{0}.$$

Since the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ are distinct, and none of the numbers $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}$ is zero, we deduce that the set of $m-1$ vectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{m-1}\}$ is linearly dependent. This, however, is impossible since we assumed that m is the *smallest* number such that a set of m vectors in E is linearly dependent. This contradiction establishes the result. ■

3 Symmetric matrices

In this section you will concentrate on diagonalising symmetric matrices. You will see that such matrices are always diagonalisable and that their transition matrices can be chosen to have particular properties.

3.1 Diagonalising symmetric matrices

Suppose that $\mathbf{A}$ is an $n \times n$ matrix and that we can find a basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ for $\mathbb{R}^n$ consisting of eigenvectors of $\mathbf{A}$. In Section 2 you saw that $\mathbf{A}$ can be diagonalised: if $\mathbf{P}$ is the transition matrix whose columns are formed from the coordinates of the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, then

$$\mathbf{P}^{-1}\mathbf{AP}$$

is a diagonal matrix.

In this section you will see that whenever $\mathbf{A}$ is an $n \times n$ *symmetric* matrix (a matrix where $\mathbf{A}^T = \mathbf{A}$), then we can always find a basis for $\mathbb{R}^n$ made up of eigenvectors of $\mathbf{A}$, and so such a matrix is always diagonalisable. In fact, we can always find an *orthonormal* basis for $\mathbb{R}^n$ made up of eigenvectors of $\mathbf{A}$. Recall from Subsection 5.4 of Unit C2 that an *orthonormal* basis

consists of mutually perpendicular (*orthogonal*) vectors of magnitude 1. For example, the standard basis for $\mathbb{R}^n$ is an orthonormal basis.

When we have an orthonormal basis, it turns out that the inverse of the transition matrix $\mathbf{P}$ is actually the transpose of $\mathbf{P}$; that is, $\mathbf{P}^{-1} = \mathbf{P}^T$. This can be useful since finding the transpose of a matrix is much simpler than finding the inverse. We will prove this result as Theorem C65 in the next subsection where you will also see that orthogonal matrices have other useful properties.

For example, consider the symmetric matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

We will show that there is an orthonormal basis for $\mathbb{R}^3$ that consists of eigenvectors of $\mathbf{A}$.

You found in Exercise C120 that the eigenvalues of $\mathbf{A}$ are $\lambda = 6$, $\lambda = 3$ and $\lambda = 0$, and that the eigenvectors are the non-zero vectors of the following forms:

$$\begin{aligned}
&(2k, 2k, k), \text{ corresponding to } \lambda = 6,\\
&(-2k, k, 2k), \text{ corresponding to } \lambda = 3,\\
&(k, -2k, 2k), \text{ corresponding to } \lambda = 0.
\end{aligned}$$

Exercise C135

Let $\mathbf{v}_1 = (2k, 2k, k)$, $\mathbf{v}_2 = (-2l, l, 2l)$ and $\mathbf{v}_3 = (m, -2m, 2m)$, where k, l, m are positive real numbers.

(a) Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an *orthogonal* basis for $\mathbb{R}^3$.

(b) Find values of k, l and m for which $|\mathbf{v}_1| = |\mathbf{v}_2| = |\mathbf{v}_3| = 1$.

In Subsection 5.4 of Unit C2 you saw that $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for $i \neq j$, and $|\mathbf{v}_i| = 1$ for each i. It follows from Exercise C135 that

$$E = \left\{ \left(\tfrac{2}{3}, \tfrac{2}{3}, \tfrac{1}{3}\right), \left(-\tfrac{2}{3}, \tfrac{1}{3}, \tfrac{2}{3}\right), \left(\tfrac{1}{3}, -\tfrac{2}{3}, \tfrac{2}{3}\right) \right\}$$

is an orthonormal basis for $\mathbb{R}^3$. Since E is an eigenvector basis of $\mathbf{A}$, we say that E is an *orthonormal eigenvector basis* of $\mathbf{A}$.

Following Strategy C20, we diagonalise the matrix $\mathbf{A}$ by writing down the transition matrix $\mathbf{P}$ whose columns are formed from the standard coordinates of the vectors in E:

$$\mathbf{P} = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}.$$

A transition matrix formed from an orthonormal eigenvector basis in this way is called an *orthogonal* matrix.

Definition

An $n \times n$ matrix whose columns form an orthonormal basis for $\mathbb{R}^n$ is an **orthogonal** matrix.

It is important to remember that the columns of an orthogonal matrix are *orthonormal* vectors, not just *orthogonal* vectors, despite the name!

Consider the 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

The columns of $\mathbf{A}$ (as vectors) are orthogonal since

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \cdot \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = 0.$$

Orthogonal vectors are linearly independent, so the columns of $\mathbf{A}$ form a basis for $\mathbb{R}^2$.

The columns of $\mathbf{A}$ (as vectors) also have magnitude 1 since

$$\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = 1, \quad \left(\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2 = 1,$$

so the matrix $\mathbf{A}$ is an orthogonal matrix.

Exercise C136

Show that $\mathbf{P}^T\mathbf{P} = \mathbf{I}$, where

$$\mathbf{P} = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}.$$

($\mathbf{P}$ is the orthogonal matrix formed below Exercise C135.)

We know that if $\mathbf{P}^T\mathbf{P} = \mathbf{I}$, then $\mathbf{P}\mathbf{P}^T = \mathbf{I}$ (by Theorem C18 in Unit C1), so for the matrix $\mathbf{P}$ in Exercise C136, $\mathbf{P}^T$ is the inverse of $\mathbf{P}$; that is, $\mathbf{P}^T = \mathbf{P}^{-1}$. We will prove that $\mathbf{P}^T = \mathbf{P}^{-1}$ for *any* orthogonal matrix $\mathbf{P}$ as Theorem C65 in the next subsection.

It follows from this and Strategy C20 that

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We say that the matrix $\mathbf{A}$ has been *orthogonally diagonalised*.

Definition

The matrix $\mathbf{A}$ is **orthogonally diagonalisable** if there exists an orthogonal matrix $\mathbf{P}$ such that the matrix

$$\mathbf{D} = \mathbf{P}^T\mathbf{AP} = \mathbf{P}^{-1}\mathbf{AP}$$

is diagonal.

The following strategy is a modification of Strategy C20 for diagonalising a matrix.

Strategy C22

To orthogonally diagonalise an $n \times n$ symmetric matrix $\mathbf{A}$:

1. find all the eigenvalues of $\mathbf{A}$
2. find an orthonormal eigenvector basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of $\mathbf{A}$
3. write down the orthogonal transition matrix $\mathbf{P}$ whose jth column is formed from the standard coordinates of $\mathbf{e}_j$.

Then

$$\mathbf{P}^T\mathbf{AP} = \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_j is the eigenvalue corresponding to the eigenvector $\mathbf{e}_j$.

In Section 4 you will see that orthogonal diagonalisation is used for classifying conics and quadrics. However, if the aim is simply to *diagonalise* a symmetric matrix as opposed to *orthogonally diagonalise* it, then use Strategy C20 – this saves time and effort when an orthonormal basis, or equivalently an orthogonal transition matrix, is not required. It is always a good idea to consider carefully what a problem requires you to do in order to solve it in the most efficient way.

You may have noticed that the words 'if possible' appear in Strategy C20, but not in Strategy C22. This is due to the fact that an $n \times n$ symmetric matrix $\mathbf{A}$ *always* has an orthonormal eigenvector basis, so it must be orthogonally diagonalisable. It is also true that any orthogonally diagonalisable matrix $\mathbf{A}$ must be symmetric – you might like to prove this yourself; it is included as a 'challenging' exercise in the additional exercises booklet for this unit.

In the case where a symmetric matrix $\mathbf{A}$ has n distinct eigenvalues, the fact that $\mathbf{A}$ has an orthonormal eigenvector basis follows from the following result.

Theorem C64

Eigenvectors corresponding to distinct eigenvalues of a symmetric matrix are orthogonal.

Proof Let $\mathbf{A}$ be a symmetric matrix, and let $\mathbf{v}$ and $\mathbf{w}$ be eigenvectors of $\mathbf{A}$ corresponding to the distinct eigenvalues λ and μ. Then

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \quad \text{and} \quad \mathbf{A}\mathbf{w} = \mu\mathbf{w}.$$

To show that $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, we need to show that $\mathbf{v} \cdot \mathbf{w} = 0$. We do this by writing $\mathbf{v}^T\mathbf{A}\mathbf{w}$ in two ways and using the fact that $\mathbf{v}^T\mathbf{w} = \mathbf{v} \cdot \mathbf{w}$. This fact is illustrated in Figure 9.

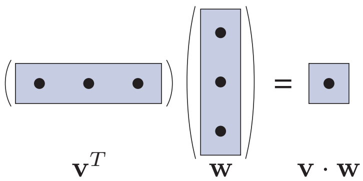


Figure 9 $\mathbf{v}^T\mathbf{w} = \mathbf{v} \cdot \mathbf{w}$

We have,

$$\mathbf{v}^T\mathbf{A}\mathbf{w} = \mathbf{v}^T(\mathbf{A}\mathbf{w}) = \mathbf{v}^T(\mu\mathbf{w}) = \mu(\mathbf{v}^T\mathbf{w}) = \mu(\mathbf{v} \cdot \mathbf{w}).$$

Since $\mathbf{A}$ is symmetric, we have $\mathbf{A}^T = \mathbf{A}$, and therefore that

$$\mathbf{v}^T\mathbf{A} = \mathbf{v}^T\mathbf{A}^T = (\mathbf{A}\mathbf{v})^T.$$

It follows that

$$\mathbf{v}^T\mathbf{A}\mathbf{w} = (\mathbf{v}^T\mathbf{A})\mathbf{w} = (\mathbf{A}\mathbf{v})^T\mathbf{w} = (\lambda\mathbf{v})^T\mathbf{w} = \lambda(\mathbf{v}^T\mathbf{w}) = \lambda(\mathbf{v} \cdot \mathbf{w}).$$

Therefore $\lambda(\mathbf{v} \cdot \mathbf{w}) = \mu(\mathbf{v} \cdot \mathbf{w})$; thus

$$(\lambda - \mu)(\mathbf{v} \cdot \mathbf{w}) = 0.$$

Since the eigenvalues λ and μ are distinct, $\lambda - \mu$ is non-zero, and hence $\mathbf{v} \cdot \mathbf{w} = 0$. The two eigenvectors $\mathbf{v}$ and $\mathbf{w}$ are orthogonal as required. ■

The following exercises show how Theorem C64 can be used.

Worked Exercise C71

Orthogonally diagonalise the symmetric matrix

$$\mathbf{A} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}.$$

Solution

We use Strategy C22.

We found the eigenvalues and eigenvectors of $\mathbf{A}$ in Worked Exercise C63.

The eigenvalues of $\mathbf{A}$ are $\lambda = 7$ and $\lambda = 3$.

The eigenvectors of $\mathbf{A}$ are the non-zero vectors of the following form:

$$(k, k), \ \text{corresponding to } \lambda = 7,$$
$$(k, -k), \ \text{corresponding to } \lambda = 3.$$

Since any eigenvectors corresponding to these eigenvalues are orthogonal by Theorem C64, we form an orthonormal eigenvector basis of $\mathbf{A}$ by taking an eigenvector of magnitude 1 corresponding to each of the two distinct eigenvalues.

An eigenvector of magnitude 1 corresponding to $\lambda = 7$ is

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

An eigenvector of magnitude 1 corresponding to $\lambda = 3$ is

$$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

It follows from Theorem C64 that an orthonormal eigenvector basis of $\mathbf{A}$ is

$$E = \left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\right\}.$$

We use the eigenvectors in E to form the columns of the orthogonal transition matrix:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^T\mathbf{AP} = \mathbf{D} = \begin{pmatrix} 7 & 0 \\ 0 & 3 \end{pmatrix}.$$

Exercise C137

Orthogonally diagonalise each of the following symmetric matrices.

(a) $\mathbf{A} = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 5 & -1 & -1 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix}$

The eigenvalues in part (b) are $\lambda = 6$, $\lambda = 3$ and $\lambda = 2$.

So far, in each case where we have orthogonally diagonalised an $n \times n$ symmetric matrix, we have had n distinct eigenvalues and Theorem C64 has ensured that the eigenvectors are all orthogonal. We have then formed an orthonormal eigenvector basis for the matrix by writing down basis vectors of magnitude 1. Where the eigenvalues of the symmetric matrix are not all distinct we have to find an orthonormal eigenvector basis for each eigenspace – then Theorem C64 will ensure that the resulting set of eigenvectors will form an orthonormal eigenvector basis for the matrix.

The following strategy is a modification of Strategy C21. It reflects the fact that we can always find an orthonormal basis comprising r vectors for an eigenspace of a *symmetric* matrix corresponding to an eigenvalue of multiplicity r. This result is not proved here.

Strategy C23

To find an orthonormal eigenvector basis of a *symmetric* matrix $\mathbf{A}$:

1. find an orthonormal basis for each eigenspace of $\mathbf{A}$
2. form the set E of all the basis vectors found in step 1.

Then E is an orthonormal eigenvector basis of $\mathbf{A}$.

Worked Exercise C72

Orthogonally diagonalise the symmetric matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}.$$

Solution

In Worked Exercise C70 we found an eigenvector basis of $\mathbf{A}$: $E = \{(1,1,1), (1,0,-1), (0,1,-1)\}$ corresponding to the eigenvalues $\lambda = 8$, $\lambda = 2$ and $\lambda = 2$, respectively.

A basis for the eigenspace $S(8)$ is $\{(1,1,1)\}$, so an orthonormal basis for $S(8)$ is

$$\left\{ \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right) \right\}.$$

A basis for the eigenspace $S(2)$ is $\{(1,0,-1), (0,1,-1)\}$.

These two basis vectors are not orthogonal since

$$(1,0,-1) \cdot (0,1,-1) = 1 \neq 0.$$

We need to find a pair of orthogonal vectors that span $S(2)$. One method for this is the Gram–Schmidt orthogonalisation process that you met in Subsection 5.3 of Unit C2.

To find an *orthogonal* basis for the eigenspace $S(2)$, we use the Gram–Schmidt orthogonalisation process.

Let the orthogonal basis we seek be $\{\mathbf{v}_1, \mathbf{v}_2\}$, with $\mathbf{v}_1 = (1, 0, -1)$.

Then

$$\begin{aligned}
\mathbf{v}_2 &= (0, 1, -1) - \left(\frac{\mathbf{v}_1 \cdot (0, 1, -1)}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 \\
&= (0, 1, -1) - \left(\frac{(1, 0, -1) \cdot (0, 1, -1)}{(1, 0, -1) \cdot (1, 0, -1)}\right)(1, 0, -1) \\
&= (0, 1, -1) - \tfrac{1}{2}(1, 0, -1) \\
&= \left(-\tfrac{1}{2}, 1, -\tfrac{1}{2}\right).
\end{aligned}$$

Dividing $\mathbf{v}_2$ by $|\mathbf{v}_2| = \sqrt{6}/2$ gives a unit basis vector. However, although it is not necessary it is often helpful to minimise the minus signs involved: we can multiply through by -1 to get another unit basis vector orthogonal to $\mathbf{v}_1$.

An orthonormal basis for $S(2)$ is therefore

$$\left\{\left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)\right\}.$$

We have ensured that the eigenvectors in the basis for $S(2)$ are orthogonal, and by Theorem C64 the eigenvectors corresponding to the distinct eigenvalues $\lambda = 8$ and $\lambda = 2$ are orthogonal.

By Theorem C64 an orthonormal eigenvector basis of $\mathbf{A}$ is therefore

$$E = \left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)\right\}.$$

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}.$$

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

The diagonal matrix found here is the same as that found in Worked Exercise C70, since the eigenvalues are considered in the same order. The difference in the diagonalisation lies in the transition matrix, which in this case is orthogonal.

Exercise C138

Orthogonally diagonalise the symmetric matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

(In Exercise C134 you found the eigenvalues and eigenvectors of $\mathbf{A}$: that a basis for $S(3)$ is $\{(0,1,1)\}$ and a basis for $S(1)$ is $\{(1,0,0),(0,1,-1)\}$.)

We conclude this subsection by noting that every symmetric matrix can be orthogonally diagonalised and conversely that an orthogonally diagonalisable matrix is symmetric. However, it is possible to diagonalise (but not orthogonally diagonalise) a non-symmetric matrix that has an eigenvector basis.

3.2 Orthogonal matrices

In this subsection we look at some properties of orthogonal matrices. Remember that the columns of an orthogonal matrix form an orthonormal basis, not merely an orthogonal basis; that is, the columns are orthogonal vectors of magnitude 1.

We have said that whenever $\mathbf{P}$ is an orthogonal matrix we have $\mathbf{P}^T = \mathbf{P}^{-1}$. We now prove this result.

Theorem C65

A square matrix $\mathbf{P}$ is orthogonal if and only if $\mathbf{P}^T = \mathbf{P}^{-1}$.

Proof We know by Theorem C18 in Unit C1 that $\mathbf{P}^T\mathbf{P} = \mathbf{I}$ if and only if $\mathbf{P}\mathbf{P}^T = \mathbf{I}$, so $\mathbf{P}^T = \mathbf{P}^{-1}$ if and only if $\mathbf{P}^T\mathbf{P} = \mathbf{I}$.

So we need to show that $\mathbf{P}$ is orthogonal if and only if $\mathbf{P}^T\mathbf{P} = \mathbf{I}$. We start off by considering the expression $\mathbf{P}^T\mathbf{P}$.

Let the columns of the matrix $\mathbf{P}$ be the column vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. Then the rows of the matrix $\mathbf{P}^T$ are the row vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.

For each i and j, the (i,j)-entry of $\mathbf{P}^T\mathbf{P}$ is the scalar product of the ith row of $\mathbf{P}^T$ and the jth column of $\mathbf{P}$; that is, $\mathbf{x}_i \cdot \mathbf{x}_j$.

So $\mathbf{P}^T\mathbf{P} = \mathbf{I}$ if and only if

$$\mathbf{x}_i \cdot \mathbf{x}_j = 0 \text{ whenever } i \neq j \quad \text{and} \quad \mathbf{x}_i \cdot \mathbf{x}_i = 1 \text{ for each } i.$$

This is the case precisely when $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is an orthonormal basis for $\mathbb{R}^n$; that is, when $\mathbf{P}$ is orthogonal. ■

Several properties of orthogonal matrices follow from Theorem C65.

Corollary C66

Let $\mathbf{P}$ and $\mathbf{Q}$ be orthogonal $n \times n$ matrices. Then:

(a) $\mathbf{P}^{-1}(=\mathbf{P}^T)$ is orthogonal

(b) the rows of $\mathbf{P}$ form an orthonormal basis for $\mathbb{R}^n$

(c) $\det \mathbf{P} = \pm 1$

(d) the product $\mathbf{PQ}$ is orthogonal.

Proof (a) To show that $\mathbf{P}^{-1}$ is orthogonal we must show that the transpose of $\mathbf{P}^{-1}$ is the inverse of $\mathbf{P}^{-1}$.

By Theorem C65 we have $\mathbf{P}^T = \mathbf{P}^{-1}$. Now,

$$(\mathbf{P}^{-1})^T\mathbf{P}^{-1} = (\mathbf{P}^T)^T\mathbf{P}^{-1} = \mathbf{P}\mathbf{P}^{-1} = \mathbf{I}.$$

Thus $(\mathbf{P}^{-1})^T = (\mathbf{P}^{-1})^{-1}$, so $\mathbf{P}^{-1}(=\mathbf{P}^T)$ is orthogonal.

(b) The rows of $\mathbf{P}$ are the columns of $\mathbf{P}^T$. The matrix $\mathbf{P}^T$ is orthogonal by part (a), so its columns form an orthonormal basis for $\mathbb{R}^n$. Thus the rows of $\mathbf{P}$ form an orthonormal basis for $\mathbb{R}^n$.

(c) We know that $\det \mathbf{P}^T = \det \mathbf{P}$, and $\mathbf{P}^T = \mathbf{P}^{-1}$ by Theorem C65, so $\mathbf{P}^T\mathbf{P} = \mathbf{I}$.

By Theorem C14 in Unit C1 we know that $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$ and $\det \mathbf{A}^T = \det \mathbf{A}$ for square matrices $\mathbf{A}$ and $\mathbf{B}$ of the same size.

Now,

$$\det(\mathbf{P}^T\mathbf{P}) = (\det \mathbf{P}^T)(\det \mathbf{P}) = (\det \mathbf{P})^2,$$

but

$$\det(\mathbf{P}^T\mathbf{P}) = \det \mathbf{I} = 1,$$

so $(\det \mathbf{P})^2 = 1$. Hence $\det \mathbf{P} = \pm 1$.

(d) The proof of this is left for you to do in Exercise C139. ■

Exercise C139

Let $\mathbf{P}$ and $\mathbf{Q}$ be orthogonal $n \times n$ matrices. Prove that the product $\mathbf{PQ}$ is orthogonal.

(This is part (d) of Corollary C66.)

To understand why orthogonal diagonalisation is useful – beyond the ease of finding the inverse of the transition matrix – we will now look at the geometry of orthogonal transition matrices in $\mathbb{R}^2$ and $\mathbb{R}^3$.

We begin by asking to what transformations of the plane the 2×2 orthogonal matrices correspond. Suppose that

$$\mathbf{P} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is an orthogonal matrix. Then the vectors (a, c) and (b, d) form an orthonormal basis for $\mathbb{R}^2$ and $\det \mathbf{P} = \pm 1$.

We stated in Subsection 3.2 of Unit C3 that the magnitude of the determinant of a matrix of a linear transformation gives the 'scaling factor'. Therefore $\det \mathbf{P} = \pm 1$ means that there is no scaling; that is, magnitudes are preserved.

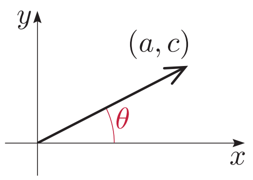


Figure 10 The angle θ made by the vector (a, c)

Let θ be the angle that the unit vector (a, c) makes with the x-axis, as illustrated in Figure 10 for the case that (a, c) is in the first quadrant, so

$$(a, c) = (\cos\theta, \sin\theta).$$

Since the unit vector (b, d) is orthogonal to (a, c), we have $(a, c) \cdot (b, d) = 0$, so

$$(b, d) = (-\sin\theta, \cos\theta) \quad \text{or} \quad (\sin\theta, -\cos\theta),$$

as illustrated in Figure 11.

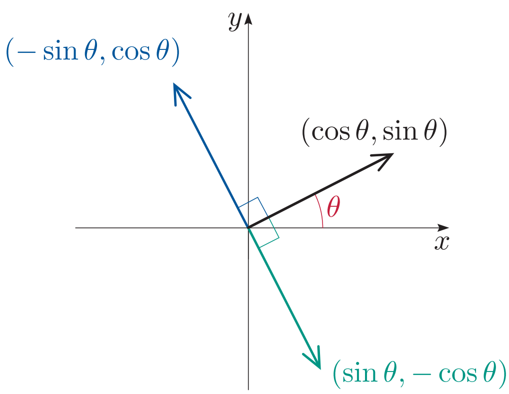


Figure 11 The two possible vectors (b, d) orthogonal to the vector (a, c)

Hence, if $\det \mathbf{P} = +1$, then

$$\mathbf{P} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

and if $\det \mathbf{P} = -1$, then

$$\mathbf{P} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

Now suppose that $E = \{\mathbf{e}_1, \mathbf{e}_2\}$ is an orthonormal basis for $\mathbb{R}^2$ and that $\mathbf{P}$ is the orthogonal transition matrix whose columns are formed from the coordinates of $\mathbf{e}_1$ and $\mathbf{e}_2$.

We have just seen that if $\det \mathbf{P} = +1$, then

$$\mathbf{e}_1 = (\cos\theta, \sin\theta) \quad \text{and} \quad \mathbf{e}_2 = (-\sin\theta, \cos\theta),$$

that is, $\mathbf{e}_1$ and $\mathbf{e}_2$ are the images of the standard basis vectors $(1, 0)$ and $(0, 1)$ under a rotation r_θ, as illustrated in Figure 12.

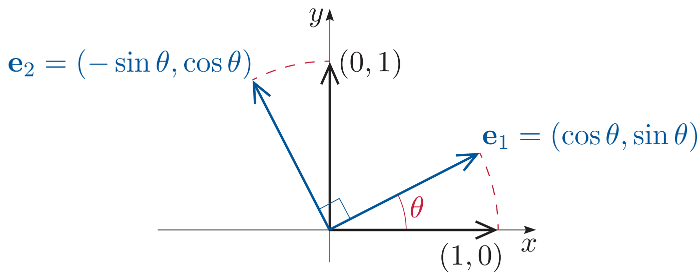


Figure 12 A rotation r_θ

Similarly, if $\det \mathbf{P} = -1$, then $\mathbf{e}_1$ and $\mathbf{e}_2$ are the images of the standard basis vectors $(1, 0)$ and $(0, 1)$ under a reflection $q_{\theta/2}$, as illustrated in Figure 13.

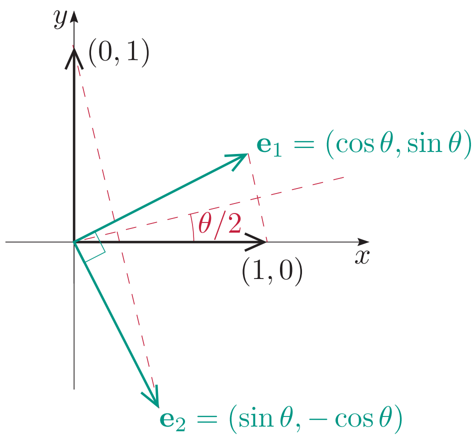


Figure 13 A reflection $q_{\theta/2}$

So if a 2×2 orthogonal matrix $\mathbf{P}$ is used to represent a linear transformation (as opposed to a transition matrix), then the linear transformation must be either a rotation or a reflection.

Similar arguments can be applied to 3×3 orthogonal matrices to show that linear transformations of $\mathbb{R}^3$ whose matrices are orthogonal are rotations about a line through the origin, reflections in a plane through the origin or combinations of these. The orthogonal matrices representing rotations of $\mathbb{R}^3$ are precisely those with determinant $+1$.

Exercise C140

Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(a) Verify that this matrix is orthogonal.

(b) Show that this matrix represents a rotation of $\mathbb{R}^3$.

Let t be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ with a matrix representation that is a symmetric matrix $\mathbf{A}$. In effect, when we orthogonally diagonalise $\mathbf{A}$, we are finding a basis for $\mathbb{R}^n$ for which

- the matrix of t is diagonal
- the basis vectors are orthogonal
- the basis vectors have magnitude 1.

For $\mathbb{R}^2$ and $\mathbb{R}^3$ this new basis is simply the standard basis rotated, reflected or, for $\mathbb{R}^3$, a combination of the two.

4 Conics and quadrics

In this section you will classify conics and quadrics using many of the techniques you have learned in this book on linear algebra, including orthogonal diagonalisation of symmetric matrices.

You revised conics in Unit A4 *Real functions, graphs and conics*.

4.1 Classifying conics

A non-degenerate conic may be a circle, an ellipse, a parabola or a hyperbola. It is said to be in **standard position** if it is positioned in the plane as follows.

- For a circle: its centre is at the origin.
- For an ellipse: its axes of symmetry are the x- and y-axes, and its largest width is along the x-axis.
- For a parabola: its axis of symmetry is the x-axis, it passes through the origin and its other points lie to the right of the origin.
- For a hyperbola: its axes of symmetry are the x- and y-axes, and it crosses the x-axis.

A circle may sometimes be considered to be a special type of ellipse, and that will be the case throughout this section.

An ellipse, a parabola and a hyperbola in standard position are illustrated in Figure 14.

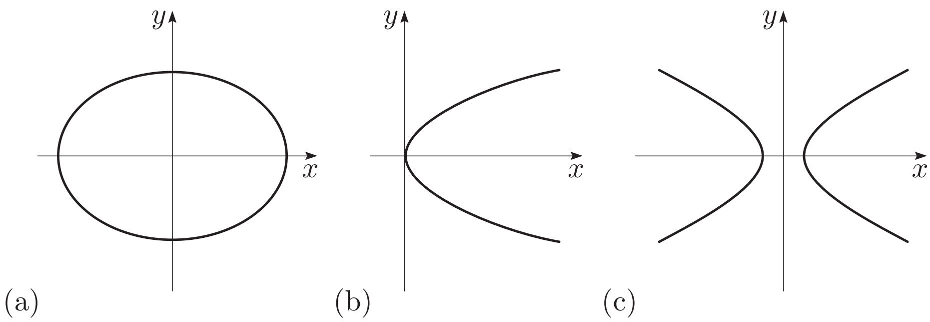


Figure 14 Conics in standard position: (a) ellipse (b) parabola and (c) hyperbola

The line joining the vertices of an ellipse is the major axis of the ellipse, and the line perpendicular to this through the centre of the ellipse is the minor axis of the ellipse. Thus, for an ellipse in standard position, the major and minor axes are the x-axis and y-axis, respectively.

We can define major and minor axes for parabolas and hyperbolas similarly.

- For a parabola, the major axis is the axis of the parabola, and the minor axis is the line perpendicular to this through the vertex of the parabola.
- For a hyperbola, the major axis is the line joining the vertices of the hyperbola, and the minor axis is the line perpendicular to this through the centre of the hyperbola.

Notice, in each case the minor axis is parallel to the directrix of the conic. (You met the directrix of a conic in Section 5 of Unit A4).

In this way, the major and minor axes of any conic in standard position are the x-axis and y-axis, respectively.

An ellipse in standard position has equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

a parabola in standard position has equation

$$y^2 = 4ax$$

and a hyperbola in standard position has equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Theorem A21 in Unit A4 says that any conic in $\mathbb{R}^2$ is the set of points (x, y) in $\mathbb{R}^2$ that satisfy an equation of the following form

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0, \tag{6}$$

where A, B, C, F, G and H are real numbers, and A, B and C are not all zero. This theorem also says the converse: that the set of all points in $\mathbb{R}^2$ whose coordinates (x, y) satisfy an equation of this form is a conic. However, such a conic may be degenerate – in this subsection we will only be concerned with non-degenerate conics.

Given the equation of a non-degenerate conic, such as

$$x^2 - 4xy - 2y^2 + 6x + 12y + 21 = 0, \tag{7}$$

we would like to be able to decide whether it represents an ellipse, a hyperbola or a parabola. We know it is not a circle because of the non-zero term in xy, but it is too complicated to easily determine more than this. Generally, the equations of conics that arise in calculations are not in standard position: thus we need some way of determining the nature of a conic from its equation.

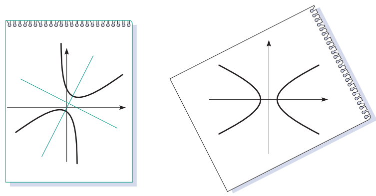


Figure 15 Moving the axes to be able to recognise a conic

In fact, equation (7) represents a hyperbola with centre $(1, 2)$, major axis $y = 2x$ and minor axis $x + 2y = 5$. This conic would be easily recognisable were we to *move the axes of the plane* so that they pass through the centre and line up with the major and minor axes of the conic, as illustrated in Figure 15.

You will see that we can move the axes of the plane by introducing matrices and changing the basis for the plane, then performing a translation so that the conic is in standard position with respect to these new basis vectors. The conic will then be easily recognisable from its equation.

We will actually be a little less specific with how we move the axes mathematically and may not always end up with a conic in standard position: the axes may be interchanged or pointing in the opposite directions resulting in conics that are reflected or rotated. However, in every case the axes will align with the major and minor axes of the conic, and the equation will resemble the equation of a conic in standard position; we say that such an equation of a conic is in **standard form**.

An ellipse and a hyperbola with equations in standard form, but that are not in standard position, are illustrated in Figure 16.

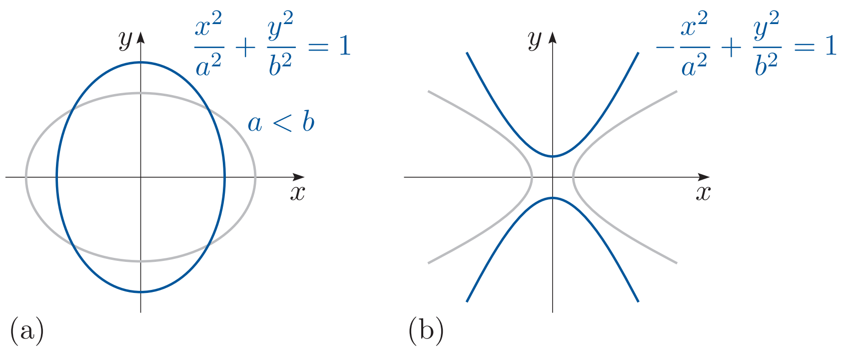


Figure 16 Conics not in standard position with equations in standard form: (a) ellipse and (b) hyperbola

Parabolas with equations in standard form, but not in standard position, are illustrated in Figure 17.

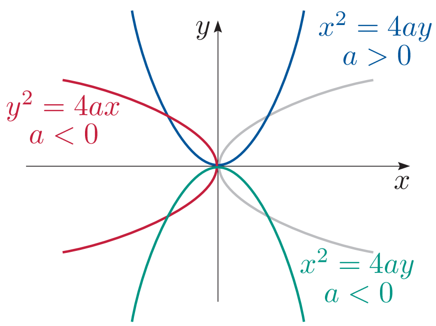


Figure 17 Parabolas not in standard position with equations in standard form

Introducing matrices

We first write equation (6) $Ax^2+Bxy+Cy^2+Fx+Gy+H=0$ using matrices and vectors; that is, in matrix form as

$$\mathbf{x}^T\mathbf{A}\mathbf{x}+\mathbf{J}^T\mathbf{x}+H=0, \tag{8}$$

where

$$\mathbf{A}=\begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix}, \quad \mathbf{J}=\begin{pmatrix} F \\ G \end{pmatrix}, \quad \mathbf{x}=\begin{pmatrix} x \\ y \end{pmatrix}.$$

This is possible, since

$$\begin{aligned}\mathbf{x}^T\mathbf{A}\mathbf{x}=\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} Ax+\frac{1}{2}By \\ \frac{1}{2}Bx+Cy \end{pmatrix} \\ &= Ax^2+Bxy+Cy^2\end{aligned}$$

and

$$\begin{aligned}\mathbf{J}^T\mathbf{x}&=\begin{pmatrix} F & G \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ &= Fx+Gy.\end{aligned}$$

Notice that the matrix $\mathbf{A}$ is symmetric; this will be important.

For example, the conic with equation (7) can be written in matrix form (8) with

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 6 \\ 12 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad H = 21.$$

Exercise C141

For each of the following equations of a conic in standard position, write the equation in matrix form and specify the matrices $\mathbf{A}$ and $\mathbf{J}$.

(a) the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ (b) the hyperbola $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$

(c) the parabola $y^2 = 4ax$

Aligning the axes

The matrix $\mathbf{A}$ in the matrix representation (8) is symmetric, so we know that we can orthogonally diagonalise this matrix to get $\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D}$ where $\mathbf{P}$ is an orthogonal transition matrix.

This helps us recognise the conic by aligning the basis vectors with the axes of the conic and therefore removing the xy-terms from the equation. The columns of $\mathbf{P}$ form an orthonormal basis E, and $\mathbf{P}$ changes E-coordinates $\mathbf{x}_E$, which we will write in the form $\mathbf{x}' = (x', y')$, into standard coordinates $\mathbf{x} = (x, y)$, so that $\mathbf{x} = \mathbf{P}\mathbf{x}'$.

In this way equation (8) becomes

$$(\mathbf{P}\mathbf{x}')^T\mathbf{A}(\mathbf{P}\mathbf{x}') + \mathbf{J}^T\mathbf{P}\mathbf{x}' + H = 0,$$

which can be rewritten as

$$(\mathbf{x}')^T(\mathbf{P}^T\mathbf{A}\mathbf{P})\mathbf{x}' + \mathbf{J}^T\mathbf{P}\mathbf{x}' + H = 0. \tag{9}$$

Now, $\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D}$ is a diagonal matrix with diagonal entries λ_1 and λ_2, so we have

$$\begin{aligned}(\mathbf{x}')^T(\mathbf{P}^T\mathbf{A}\mathbf{P})\mathbf{x}' &= (\mathbf{x}')^T\mathbf{D}\mathbf{x}' \\ &= \begin{pmatrix} x' & y' \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} \\ &= \lambda_1(x')^2 + \lambda_2(y')^2,\end{aligned}$$

and therefore there is no $x'y'$-term in the new equation (9) for the conic. Written in the form of equation (9), this now more closely resembles the equation of a conic in standard position. The vectors in the orthonormal basis E of the plane are aligned with the axes of the conic: we say we have *aligned the axes*.

The order and direction in which the eigenvectors are chosen affects the orthonormal basis E and therefore the transition matrix $\mathbf{P}$ obtained.

However, in every case $\mathbf{P}$ is an orthogonal matrix and so $\det \mathbf{P} = \pm 1$. Orthogonal diagonalisation ensures that the new basis vectors are orthogonal (perpendicular) and of magnitude 1. If $\mathbf{P}$ is considered to represent a linear transformation (as opposed to a transition matrix), then the linear transformation is either a rotation ($\det \mathbf{P} = +1$) or a reflection ($\det \mathbf{P} = -1$).

It is sometimes preferable, when choosing the orthonormal basis E, for it to be a rotation (rather than a reflection) of the standard basis vectors; that is, that $\mathbf{P}$, considered as a linear transformation, is a rotation. This is achieved by ensuring that $\det \mathbf{P} = +1$ (using either geometric insight, or by checking the determinant). However, this step is not required in this module.

We now illustrate the process of rewriting a conic in the form of equation (9) by applying the process to equation (7), where

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}.$$

Worked Exercise C73

Express the non-degenerate conic

$$x^2 - 4xy - 2y^2 + 6x + 12y + 21 = 0$$

in the form of equation (9).

Solution

The matrix form of the equation of the conic is $\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{J}^T\mathbf{x} + H = 0$ where

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 6 \\ 12 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad H = 21.$$

In Exercise C119(b) you found that the eigenvectors of $\mathbf{A}$ are the non-zero vectors $(k, 2k)$ and $(-2k, k)$, corresponding to the eigenvalues $\lambda = -3$ and $\lambda = 2$, respectively.

We start by orthogonally diagonalising $\mathbf{A}$.

We use Strategy C22 to orthogonally diagonalise $\mathbf{A}$.

An orthonormal basis for $S(-3)$ is

$$\left\{ \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right) \right\},$$

and an orthonormal basis for $S(2)$ is

$$\left\{ \left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right) \right\}.$$

By Theorem C64 an orthonormal eigenvector basis of $\mathbf{A}$ is therefore

$$E = \left\{ \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right), \left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right) \right\}.$$

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}.$$

Note that $\det \mathbf{P} = +1$, so the basis vectors in E are the images of the standard basis vectors under a rotation, but that does not concern us here.

We use the eigenvalues to form the diagonal matrix

$$\mathbf{P}^T\mathbf{AP} = \mathbf{D} = \begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix}.$$

We substitute into $(\mathbf{x}')^T(\mathbf{P}^T\mathbf{AP})\mathbf{x}' + \mathbf{J}^T\mathbf{P}\mathbf{x}' + H = 0$.

It follows from equation (9) that the equation of the conic is now

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 6 & 12 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + 21 = 0,$$

that is,

$$-3(x')^2 + 2(y')^2 + 6\sqrt{5}x' + 21 = 0.$$

There are no terms in $x'y'$ in this new equation.

You might wonder what the equation in Worked Exercise C73 would have been if the eigenvalues had been chosen in the opposite order? The next exercise investigates this.

Exercise C142

Express the non-degenerate conic

$$x^2 - 4xy - 2y^2 + 6x + 12y + 21 = 0$$

in the form of equation (9), using the eigenvalues in the order $\lambda = 2$ then $\lambda = -3$.

The equation of the conic with the eigenvalues $\lambda = -3$ then $\lambda = 2$ and the equation of the conic with the eigenvalues $\lambda = 2$ then $\lambda = -3$ are very similar. It looks like the roles of x' and y' have been interchanged; that is, the order of the coordinates have been interchanged, which corresponds to interchanging the axes. We have $\det \mathbf{P} = -1$ in Exercise C142 so this transition matrix corresponds to a reflection of the axes, whereas we have $\det \mathbf{P} = 1$ in Worked Exercise C73 so this transition matrix corresponds to a rotation.

In general, for any conic, if

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

then equation (9) is of the form

$$\lambda_1(x')^2 + \lambda_2(y')^2 + fx' + gy' + H = 0, \tag{10}$$

where $\begin{pmatrix} f & g \end{pmatrix} = \mathbf{J}^T\mathbf{P}$.

The equation of the conic in this form has been simplified since it now has no $x'y'$ terms, but is not yet in a form from which we can easily recognise the type of the conic: a translation of the axes is also required.

Translating the origin

To write the equation of the conic in standard form from which we can easily recognise the type of the conic, we need to eliminate any superfluous linear x' and y' terms. This is achieved by translating the origin using an (α, β)-translation and moving to new coordinates $\mathbf{x}'' = (x'', y'')$: we say we have *translated the origin*.

To do this, we first *complete the squares* in the equation of the conic. We illustrate this process using the conic with equation (7). We have already aligned the axes to obtain the equation

$$-3(x')^2 + 2(y')^2 + 6\sqrt{5}x' + 21 = 0,$$

which is equivalent to

$$-3\left((x')^2 - 2\sqrt{5}x'\right) + 2(y')^2 + 21 = 0.$$

This equation has no linear y' term, so we only need to complete the square involving x'. We obtain

$$-3(x' - \sqrt{5})^2 + 15 + 2(y')^2 + 21 = 0,$$

so

$$-3(x' - \sqrt{5})^2 + 2(y')^2 + 36 = 0.$$

In Subsection 1.3 of Unit A4 you saw that applying an (α, β)-translation to the graph of $y = f(x)$ gives the graph of $y = f(x - \alpha) + \beta$, or equivalently, $y - \beta = f(x - \alpha)$. We can express this translated curve more simply by using new (x', y')-coordinates obtained by an (α, β)-translation of the (x, y)-axes: we do this by setting $x' = x - \alpha$ and $y' = y - \beta$. In this new (x', y')-coordinate system the equation of the translated curve is $y' = f(x')$.

For our conic we use a $(\sqrt{5}, 0)$-translation, so we set the new coordinates to be

$$\mathbf{x}'' = (x'', y'') = (x' - \sqrt{5}, y').$$

Thus we rewrite the equation of the conic using these coordinates by substituting

$$x'' = x' - \sqrt{5} \quad \text{and} \quad y'' = y',$$

which results in the following simplified equation of the conic

$$-3(x'')^2 + 2(y'')^2 = -36,$$

or

$$\frac{(x'')^2}{12} - \frac{(y'')^2}{18} = 1.$$

This equation is now recognisable as the equation of a hyperbola in standard form. In fact, it is also a hyperbola in standard position with respect to these new axes, since the $(x'')^2$ term is positive and the $(y'')^2$ term is negative.

For this conic we have

- introduced matrices $\mathbf{A}$ and $\mathbf{J}$
- orthogonally diagonalised the matrix $\mathbf{A}$ to find the orthogonal transition matrix $\mathbf{P}$ which rotates the (x, y)-axes by $\theta = \cos^{-1}(1/\sqrt{5})$ to get the (x', y')-axes
- translated by $\sqrt{5}$ in the x' direction to get the (x'', y'')-axes.

This is illustrated in Figure 18.

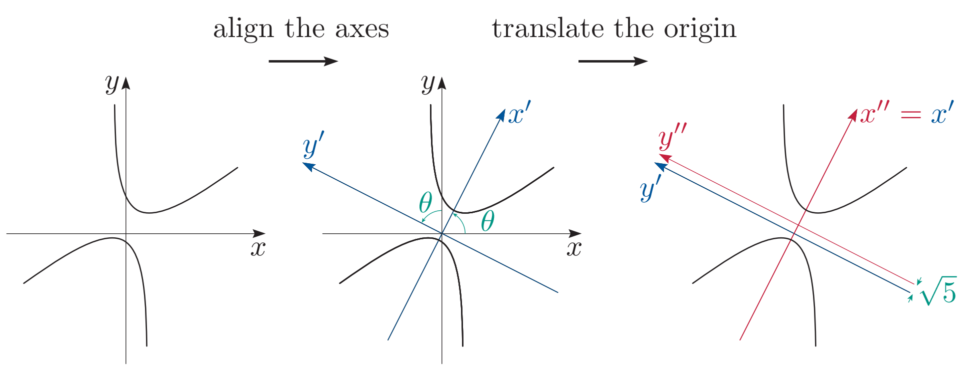


Figure 18 Moving the axes to get the equation of the conic in standard form ($\lambda = -3$ then $\lambda = 2$)

What would the equation of this conic have been if the eigenvalues had been chosen in the opposite order? The next exercise investigates this using the equation you found in Exercise C142.

Exercise C143

Write the equation of the conic

$$x^2 - 4xy - 2y^2 + 6x + 12y + 21 = 0$$

in standard form by completing the square in the equation

$$2(x')^2 - 3(y')^2 + 6\sqrt{5}y' + 21 = 0$$

and then making a substitution to get coordinates (x'', y'').

Figure 19 illustrates how the axes have been moved with the eigenvalues in the order $\lambda = 2$ then $\lambda = -3$, as in Exercise C143: the axes are reflected and then translated.

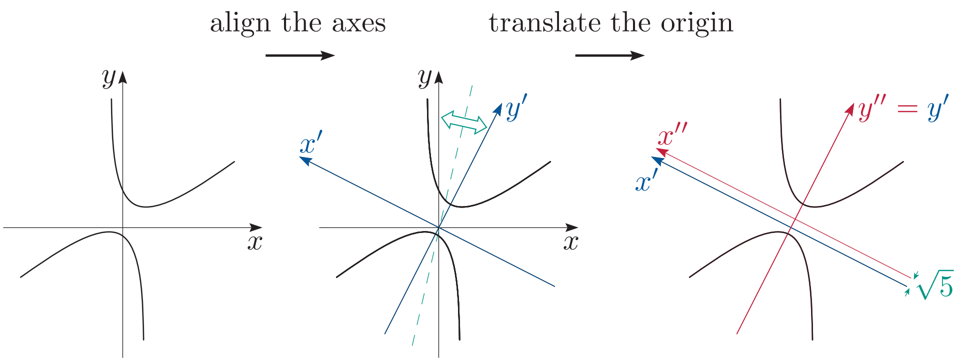


Figure 19 Moving the axes to get the equation of the conic in standard form ($\lambda = 2$ then $\lambda = -3$)

The equations in standard form found for the conic with equation (7) are

$$\frac{(x'')^2}{12} - \frac{(y'')^2}{18} = 1, \quad \text{for } \lambda = -3 \text{ then } \lambda = 2,$$

and

$$-\frac{(x'')^2}{18} + \frac{(y'')^2}{12} = 1, \quad \text{for } \lambda = 2 \text{ then } \lambda = -3.$$

In the second case the hyperbola is not in standard position with respect to these new axes, since the $(x'')^2$ term is negative and the $(y'')^2$ term is positive.

It is clear that the roles of x'' and y'' have been interchanged. Geometrically, the new axes of the plane have been interchanged, so the hyperbola has related, but different, equations in relation to these different choices of axes. However, both equations are in the standard form for a hyperbola, so the choice of the order of the eigenvalues does not affect the conclusion that this conic is a hyperbola.

Ellipse and hyperbola

In general, if neither eigenvalue is 0, then completing the squares in equation (10) gives an equation of the form

$$\lambda_1\left(x'+\frac{f}{2\lambda_1}\right)^2-\lambda_1\left(\frac{f}{2\lambda_1}\right)^2+\lambda_2\left(y'+\frac{g}{2\lambda_2}\right)^2-\lambda_2\left(\frac{g}{2\lambda_2}\right)^2+H=0,$$

which can be written as

$$\lambda_1(x'')^2+\lambda_2(y'')^2=K,$$

where

$$x''=x'+\frac{f}{2\lambda_1},\quad y''=y'+\frac{g}{2\lambda_2}\quad\text{and}\quad K=\frac{f^2}{4\lambda_1}+\frac{g^2}{4\lambda_2}-H.$$

Writing the equation in standard form gives

$$\frac{(x'')^2}{K/\lambda_1}+\frac{(y'')^2}{K/\lambda_2}=1,$$

which is the equation of an ellipse if both K/λ_1 and K/λ_2 are positive, and a hyperbola if one is negative and the other positive. (No other possibility can occur, although we do not explicitly show this.)

Parabola

In general, if one eigenvalue is 0, say λ_1 is 0 and $\lambda_2\neq 0$, then equation (10) has the form

$$\lambda_2(y')^2+fx'+gy'+H=0.$$

Completing the square in this equation gives

$$fx'+\lambda_2\left(y'+\frac{g}{2\lambda_2}\right)^2-\lambda_2\left(\frac{g}{2\lambda_2}\right)^2+H=0,$$

which can be written as

$$\lambda_2(y'')^2+fx''=0,$$

where

$$y''=y'+\frac{g}{2\lambda_2},\quad x''=x'-\frac{\lambda_2}{f}\left(\frac{g}{2\lambda_2}\right)^2+\frac{H}{f}.$$

Writing the equation in standard form gives

$$(y'')^2=\frac{-f}{\lambda_2}x'',$$

which is the equation of a parabola.

If $\lambda_1\neq 0$ and λ_2 is 0, then we obtain the similar equation

$$(x'')^2=\frac{-g}{\lambda_1}y'',$$

which is also the equation of a parabola.

Summarising the method

There are several steps involved in writing the equation of a conic in standard form, so we summarise this method in the following strategy.

Strategy C24

To write the non-degenerate conic with equation

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0$$

in standard form, do the following.

1. Introduce matrices:
   - write down $\mathbf{A} = \begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix}$ and $\mathbf{J} = \begin{pmatrix} F \\ G \end{pmatrix}$.
2. Align the axes:
   - orthogonally diagonalise $\mathbf{A}$ to get
   $$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$
   - find $\begin{pmatrix} f & g \end{pmatrix} = \mathbf{J}^T\mathbf{P}$, and write the conic in the form
   $$\lambda_1(x')^2 + \lambda_2(y')^2 + fx' + gy' + H = 0.$$
3. Translate the origin:
   - complete the squares
   - make a substitution to change to the coordinate system (x'', y'').

The order in which the eigenvalues are chosen does not affect the *form* of the equation obtained: it will be the standard form for an ellipse, a hyperbola or a parabola.

The following worked exercise and exercises illustrate this strategy.

Worked Exercise C74

Use Strategy C24 to write the non-degenerate conic with equation

$$5x^2 + 4xy + 5y^2 + 20x + 8y - 1 = 0$$

in standard form. Is this conic an ellipse, a parabola or a hyperbola?

Solution

Since some parts of this working can be quite long, we number the strategy steps in the solution.

1. Introduce matrices.
   We have
   $$\mathbf{A} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix} \quad \text{and} \quad \mathbf{J} = \begin{pmatrix} 20 \\ 8 \end{pmatrix}.$$

2. Align the axes.

   We orthogonally diagonalised $\mathbf{A}$ in Worked Exercise C71.

   We have

   $$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} 7 & 0 \\ 0 & 3 \end{pmatrix},$$

   where

   $$\mathbf{P} = \begin{pmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{pmatrix},$$

   so

   $$\begin{aligned} \begin{pmatrix} f & g \end{pmatrix} &= \begin{pmatrix} 20 & 8 \end{pmatrix} \begin{pmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{28}{\sqrt{2}} & \dfrac{12}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 14\sqrt{2} & 6\sqrt{2} \end{pmatrix}. \end{aligned}$$

   The equation of the conic is now

   $$7(x')^2 + 3(y')^2 + 14\sqrt{2}x' + 6\sqrt{2}y' - 1 = 0.$$

3. Translate the origin.

   To keep track of the terms when completing the square, we first collect the x' terms and the y' terms. We take out the coefficients of $(x')^2$ and $(y')^2$ as factors.

   We write this equation as

   $$7\left((x')^2 + 2\sqrt{2}x'\right) + 3\left((y')^2 + 2\sqrt{2}y'\right) - 1 = 0.$$

   Completing the squares in this equation, we obtain

   $$7(x' + \sqrt{2})^2 - 14 + 3(y' + \sqrt{2})^2 - 6 - 1 = 0.$$

   We substitute $x'' = x' + \sqrt{2}$ and $y'' = y' + \sqrt{2}$ into this equation and simplify to obtain

   $$7(x'')^2 + 3(y'')^2 - 21 = 0.$$

The equation of the conic in standard form is

$$\frac{(x'')^2}{3} + \frac{(y'')^2}{7} = 1.$$

The conic is an ellipse.

We can see that this ellipse is not in standard position with respect to these new axes since $3 < 7$.

Exercise C144

Use Strategy C24 to write the non-degenerate conic with equation

$$9x^2 - 4xy + 6y^2 - 10x - 20y - 5 = 0$$

in standard form. Is the conic an ellipse, a parabola or a hyperbola?

(In Exercise C137(a) you found that

$$E = \left\{\left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right), \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)\right\}$$

is an orthonormal eigenvector basis for the matrix $\mathbf{A}$ of this conic with respect to the eigenvalues $\lambda = 10$ and $\lambda = 5$.)

Exercise C145

Use Strategy C24 to write the non-degenerate conic with equation

$$x^2 - 4xy + 4y^2 - 6x - 8y + 5 = 0$$

in standard form. Is the conic an ellipse, a parabola or a hyperbola?

4.2 Classifying quadrics

Quadrics, or *quadric surfaces*, are surfaces in $\mathbb{R}^3$. They are the three-dimensional analogues of conics.

Definition

A **quadric** in $\mathbb{R}^3$ is the set of points (x, y, z) that satisfy an equation of the form

$$Ax^2 + By^2 + Cz^2 + Fxy + Gyz + Hxz + Jx + Ky + Lz + M = 0,$$

where A to M are real numbers, and A, B, C, F, G and H are not all 0.

In general the situation is more complicated than for conics and the general situation is beyond the scope of this module. However, it can be shown that there are *nine* types of quadrics involving curved surfaces in $\mathbb{R}^3$. Each of these types can be positioned in space to be in **standard position**; that is, with its axes aligned with the x-, y- and z-axes in a similar manner to the non-degenerate conics. These quadrics in standard position have easily recognisable equations and the different types can be distinguished by the **curves of intersection** of the planes parallel to the coordinate planes that meet the quadric in a non-trivial intersection. Figure 20 shows some curves of intersection for a sphere – they are all circles.

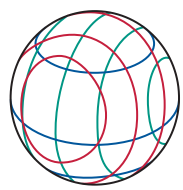


Figure 20 Some curves of intersection of a sphere

The curves of intersection of a **non-degenerate quadric** are non-degenerate conics. There are five types of non-degenerate quadric:

- the *ellipsoid* (which includes the sphere)
- the *elliptic paraboloid*
- the *hyperbolic paraboloid*
- the *hyperboloid of one sheet*
- the *hyperboloid of two sheets*.

Table 1 illustrates each of these quadrics and gives the equation in standard position, as well as specifying the curves of intersection.

There are four types of **degenerate quadric** involving curved surfaces:

- the *elliptic cone*
- the *elliptic cylinder*
- the *parabolic cylinder*
- the *hyperbolic cylinder*.

The curves of intersection of these include non-degenerate conics, degenerate conics and pairs of parallel lines. The elliptic cone in standard position is illustrated in Table 1, where the equation is given and the curves of intersection specified. The elliptic cone can be considered as intermediate between the hyperboloids of one and two sheets – where the two sheets touch at a point. The three types of cylinder in standard position, illustrated in Figure 21, are surfaces whose equations do not involve z explicitly.

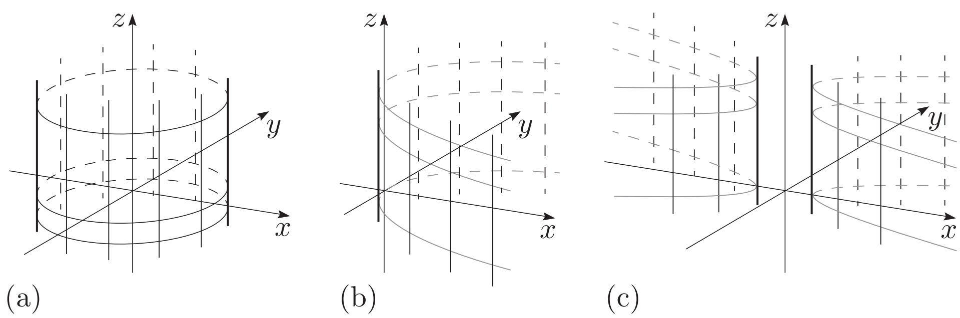


Figure 21 Degenerate quadrics: (a) elliptic cylinder (b) parabolic cylinder and (c) hyperbolic cylinder

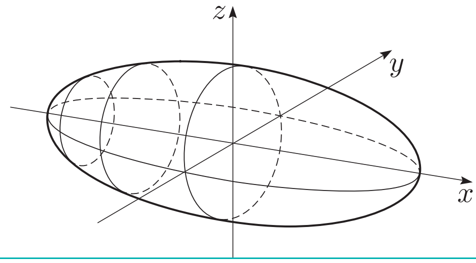
The only degenerate quadrics we will consider for the remainder of the linear algebra topic are elliptic cones, thus giving the following list of *six quadrics*, all included in Table 1: the ellipsoid (including the sphere), the elliptic paraboloid, the hyperbolic paraboloid, the hyperboloid of one sheet, the hyperboloid of two sheets and the elliptic cone.

Table 1 Quadrics: equation in standard position and the curves of intersection

Ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

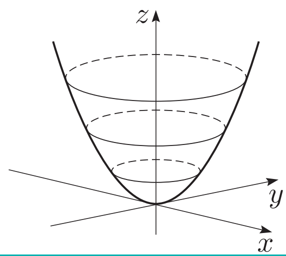
curves of intersection:
ellipse



Elliptic paraboloid

$$z=\frac{x^2}{a^2}+\frac{y^2}{b^2}$$

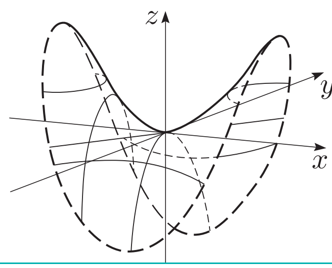
curves of intersection:
ellipse or parabola



Hyperbolic paraboloid

$$z=\frac{x^2}{a^2}-\frac{y^2}{b^2}$$

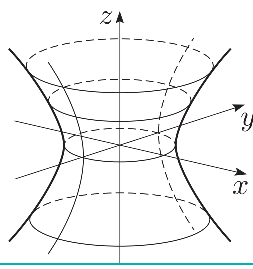
curves of intersection:
hyperbola or parabola



Hyperboloid of one sheet

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

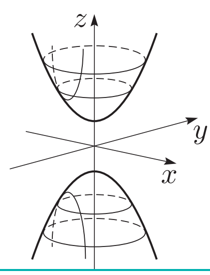
curves of intersection:
ellipse or hyperbola



Hyperboloid of two sheets

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1$$

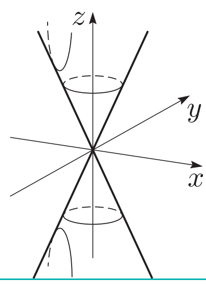
curves of intersection:
ellipse or hyperbola



Elliptic cone

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0$$

curves of intersection:
ellipse or hyperbola
(or a degenerate conic)



Gaspard Monge

The first systematic classification of quadric surfaces was by Leonhard Euler (1707–1783) in his celebrated *Introductio in analysin infinitorum* (1748) – the textbook in which he laid down the foundations of analysis – where he treated surfaces of second degree as a family of quadrics in space analogous to the plane conic sections. The subject was developed in a more rigorous way by Gaspard Monge (1746–1818) and Jean Nicolas Pierre Hachette (1769–1834) who, in 1802, provided an algebraic study of quadric surfaces, which was later published as a textbook. Both Monge and Hachette were professors at the famous École Polytechnique in Paris. This college was founded at the end of the nineteenth century to provide students with a mathematical and scientific education, and to prepare them for entry to the prestigious Grandes Écoles, higher education establishments for the training of civil and military engineers.

Jean Nicolas Pierre Hachette

As with conics, to identify a given quadric from its equation, we will align the axes and translate the origin to obtain an equation that resembles the equation of a quadric in standard position: we say that such an equation of a quadric is in **standard form**. So, for example, the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

is an equation of a hyperboloid of one sheet in standard form, although it is not in standard position.

To write the equation of a quadric in standard form, we use the same techniques that we used for conics: introducing matrices, orthogonal diagonalisation and completing the square. We omit the justification – it is analogous to that for conics.

We summarise this method in the following strategy.

Strategy C25

To write the quadric with equation

$$Ax^2 + By^2 + Cz^2 + Fxy + Gyz + Hxz + Jx + Ky + Lz + M = 0$$

in standard form, do the following.

1. Introduce matrices:
   - write down the matrices

$$\mathbf{A} = \begin{pmatrix} A & \frac{1}{2}F & \frac{1}{2}H \\ \frac{1}{2}F & B & \frac{1}{2}G \\ \frac{1}{2}H & \frac{1}{2}G & C \end{pmatrix} \quad \text{and} \quad \mathbf{J} = \begin{pmatrix} J \\ K \\ L \end{pmatrix}.$$

2. Align the axes:
   - orthogonally diagonalise $\mathbf{A}$ to get
   $$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$
   - find $\begin{pmatrix} f & g & h \end{pmatrix} = \mathbf{J}^T\mathbf{P}$, and write the quadric in the form $\lambda_1(x')^2 + \lambda_2(y')^2 + \lambda_3(z')^2 + fx' + gy' + hz' + M = 0$.
3. Translate the origin:
   - complete the squares
   - make a substitution to change to the coordinate system (x'', y'', z'').

The following worked exercise and exercises illustrate this strategy.

Worked Exercise C75

Use Strategy C25 to write the quadric with equation
$$5x^2 + 3y^2 + 3z^2 - 2xy + 2yz - 2xz - 10x + 6y - 2z - 9 = 0$$
in standard form. Which of the six types of quadric does this represent?

Solution

As with conics, since some parts of this working can be quite long, we number the strategy steps in the solution.

1. Introduce matrices.

   We have
   $$\mathbf{A} = \begin{pmatrix} 5 & -1 & -1 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{J} = \begin{pmatrix} -10 \\ 6 \\ -2 \end{pmatrix}.$$
2. Align the axes.

   You orthogonally diagonalised $\mathbf{A}$ in Exercise C137(b).

   We have
   $$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$
   where
   $$\mathbf{P} = \begin{pmatrix} -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Since $\det \mathbf{P} = 1$, this transition matrix represents a rotation of the basis vectors, but this fact does not concern us here.

So

$$\begin{aligned}\begin{pmatrix} f & g & h \end{pmatrix} &= \begin{pmatrix} -10 & 6 & -2 \end{pmatrix} \begin{pmatrix} -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 4\sqrt{6} & -2\sqrt{3} & 4\sqrt{2} \end{pmatrix}.\end{aligned}$$

The equation of the quadric is now

$$6(x')^2 + 3(y')^2 + 2(z')^2 + 4\sqrt{6}x' - 2\sqrt{3}y' + 4\sqrt{2}z' - 9 = 0.$$

3. Translate the origin.

We write this equation as

$$6\left((x')^2 + \frac{4}{\sqrt{6}}x'\right) + 3\left((y')^2 - \frac{2}{\sqrt{3}}y'\right) + 2\left((z')^2 + 2\sqrt{2}z'\right) - 9 = 0.$$

Completing the squares in this equation, we obtain

$$6\left(x' + \frac{2}{\sqrt{6}}\right)^2 - 4 + 3\left(y' - \frac{1}{\sqrt{3}}\right)^2 - 1 + 2(z' + \sqrt{2})^2 - 4 - 9 = 0.$$

Substituting

$$x'' = x' + \frac{2}{\sqrt{6}}, \quad y'' = y' - \frac{1}{\sqrt{3}} \quad \text{and} \quad z'' = z' + \sqrt{2}$$

in this equation and simplifying, we obtain

$$6(x'')^2 + 3(y'')^2 + 2(z'')^2 - 18 = 0.$$

The equation of the quadric in standard form is

$$\frac{(x'')^2}{3} + \frac{(y'')^2}{6} + \frac{(z'')^2}{9} = 1.$$

This is the equation of an ellipsoid.

Exercise C146

Use Strategy C25 to write the quadric with equation

$$x^2 + y^2 + z^2 - 2x + 4y - 6z - 11 = 0$$

in standard form. Which of the six types of quadric does this represent?

Exercise C147

Use Strategy C25 to write the quadric with equation

$$4x^2 + 3y^2 + 2z^2 + 4xy + 4yz + 12x + 12z + 18 = 0$$

in standard form. Which of the six types of quadric does this represent?

(At the start of Subsection 3.1 we found that

$$E = \left\{\left(\tfrac{2}{3}, \tfrac{2}{3}, \tfrac{1}{3}\right), \left(-\tfrac{2}{3}, \tfrac{1}{3}, \tfrac{2}{3}\right), \left(\tfrac{1}{3}, -\tfrac{2}{3}, \tfrac{2}{3}\right)\right\}$$

is an orthonormal eigenvector basis for the matrix $\mathbf{A}$ of this quadric with respect to the eigenvalues $\lambda = 6$, $\lambda = 3$ and $\lambda = 0$.)

Summary

In this unit you have met eigenvectors and eigenvalues: an eigenvector of a linear transformation $t : V \to V$ is a non-zero vector $\mathbf{v}$ that is mapped by t to a scalar multiple of itself, and this scalar is the corresponding eigenvalue λ. Since such a linear transformation always has a square matrix representation, you have seen that eigenvectors and eigenvalues can equivalently be defined in terms of matrices: $\mathbf{Av} = \lambda\mathbf{v}$. You have found eigenvalues and eigenvectors by solving the corresponding characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. You have seen that there may be no eigenvalues, for example when t is a rotation of $\mathbb{R}^2$, and that all the eigenvectors corresponding to a given eigenvalue λ, plus the zero vector, form a subspace $S(\lambda)$ of V whose dimension is never greater than the multiplicity of the eigenvalue.

You have investigated when t has an eigenvector basis E; that is, a basis comprising only eigenvectors of t, and you have met transition matrices $\mathbf{P}$ that map a basis E of V to the standard basis. You have seen (Theorem C60) that the transition matrix $\mathbf{P}$ maps standard coordinates of V to E-coordinates of V and that $\mathbf{P}$ is invertible. You have learned (Theorem C62) that whenever an eigenvector basis can be found, the transition matrix $\mathbf{P}$ can be used to express the matrix $\mathbf{A}$ of t (with respect to the standard basis) as a diagonal matrix (with respect to this eigenvector basis) via the relation $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$. Furthermore, when t has a symmetric matrix representation, the eigenvectors corresponding to different eigenvalues are orthogonal (Theorem C64), and an eigenvector basis can always be found. In addition, the basis vectors can be chosen to give an orthonormal eigenvector basis so that the transition matrix is an orthogonal matrix satisfying $\mathbf{P}^T = \mathbf{P}^{-1}$, giving $\mathbf{D} = \mathbf{P}^T\mathbf{AP}$.

Thus diagonalising matrices involves the main ideas you have studied throughout this book on linear algebra: vectors, matrices, vector spaces, bases and linear transformations.

In the final section you have seen how these techniques can be used to identify the type of a conic, or quadric, from its equation.

Learning outcomes

After working through this unit, you should be able to:

- explain the meaning of the terms *eigenvalue*, *eigenvector*, *characteristic equation* and *eigenspace*
- recognise the geometric interpretation of eigenvectors and eigenspaces in special cases
- find the eigenvalues and eigenvectors of a given 2×2 or 3×3 matrix
- describe some basic properties of eigenvalues and eigenvectors
- write down the matrix of a linear transformation t with respect to a given eigenvector basis of t
- write down the *transition matrix* from an eigenvector basis to the standard basis
- *diagonalise* a given square matrix, if possible
- understand that any symmetric matrix can be *orthogonally diagonalised*
- orthogonally diagonalise a given symmetric matrix
- describe some basic properties of *orthogonal matrices*
- write the equation of a given non-degenerate conic in standard form and hence classify it
- understand the term *quadric* and recognise the six types of quadric covered
- write the equation of a given quadric in standard form and hence classify it.

Solutions to exercises

Solution to Exercise C115

We have

$$t(2,-2) = (2-8, 2+4) = (-6,6) \\ = -3(2,-2)$$

and

$$t(-7,7) = (-7+28, -7-14) = (21,-21) \\ = -3(-7,7).$$

In each case the original vector is scaled by the factor -3.

Solution to Exercise C116

(a) We have $t(0,1) = (4,-2)$, $t(1,2) = (9,-3)$ and $t(4,1) = (8,2)$.

(b) The linear transformation t maps the line joining the points $(0,0)$ and $(4,1)$ to the line joining the points $(0,0)$ and $(8,2)$. But $(8,2) = 2(4,1)$, so these lines are the same and both can be written as $x = 4y$. Therefore the line $x = 4y$ is mapped to itself by the linear transformation t.

(c) We have

$$t(4k,k) = (4k+4k, 4k-2k) = (8k,2k) \\ = 2(4k,k),$$

so any vector lying along the line $x = 4y$ is scaled by the factor 2.

Solution to Exercise C117

(a) A reflection t in the line $y = x$ maps the point (x,y) to the point (y,x). Each point on the line $y = x$ is mapped to itself, since

$$t(k,k) = (k,k) = 1(k,k),$$

so the non-zero vectors (k,k) are eigenvectors with corresponding eigenvalue 1.

Each point on the line $y = -x$ is mapped to another point on the line $y = -x$, since

$$t(k,-k) = (-k,k) = -1(k,-k),$$

so the non-zero vectors $(k,-k)$ are eigenvectors with corresponding eigenvalue -1.

(b) A 2-dilation t maps the point (x,y) to the point $(2x,2y)$. Every line through the origin is mapped to itself; that is, every non-zero vector in the plane is an eigenvector of t. Let k and l be real numbers which are not both zero. Then

$$t(k,l) = (2k,2l) = 2(k,l),$$

so the non-zero vectors (k,l) are eigenvectors with corresponding eigenvalue 2.

(c) An anticlockwise rotation t through $\pi/2$ maps the point (x,y) to the point $(-y,x)$. No line through the origin is mapped to itself by t, so t has no eigenvectors.

(d) An anticlockwise rotation t through π maps the point (x,y) to the point $(-x,-y)$. Each line through the origin is mapped to itself; that is, each non-zero vector in the plane is an eigenvector of t. Let k and l be real numbers that are not both zero. Then

$$t(k,l) = (-k,-l) = -1(k,l),$$

so the non-zero vectors (k,l) are eigenvectors with corresponding eigenvalue -1.

Solution to Exercise C118

(a) We wish to find those vectors (x,y) that are mapped to scalar multiples of themselves; that is, the vectors that satisfy

$$(-5x+3y, 6x-2y) = (\lambda x, \lambda y).$$

Equating coordinates, we obtain the system

$$\begin{aligned} -5x + 3y &= \lambda x \\ 6x - 2y &= \lambda y, \end{aligned}$$

which we write as

$$\begin{aligned} (-5-\lambda)x + \quad\quad 3y &= 0 \\ 6x + (-2-\lambda)y &= 0. \end{aligned}$$

(b) Non-zero solutions to the eigenvector equations exist if and only if the determinant of the coefficient matrix is 0; that is, if and only if

$$\begin{vmatrix} -5-\lambda & 3 \\ 6 & -2-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(-5-\lambda)(-2-\lambda)-18=0,$$

which simplifies to

$$\lambda^2+7\lambda-8=0.$$

The eigenvalues of t are the solutions to this characteristic equation. We have

$$\lambda^2+7\lambda-8=(\lambda-1)(\lambda+8)=0,$$

so the eigenvalues are $\lambda=1$ and $\lambda=-8$.

(c) To find the corresponding eigenvectors, we consider each value of λ in turn.

$\boxed{\lambda=1}$ The eigenvector equations become

$$\begin{aligned}-6x+3y&=0\\ 6x-3y&=0.\end{aligned}$$

These equations are equivalent to the single equation

$$2x-y=0.$$

Thus the eigenvectors corresponding to $\lambda=1$ are the non-zero vectors (x,y) for which $y=2x$; that is, the vectors of the form

$$(k,2k),\quad \text{where } k\neq 0.$$

$\boxed{\lambda=-8}$ The eigenvector equations become

$$\begin{aligned}3x+3y&=0\\ 6x+6y&=0.\end{aligned}$$

These equations are equivalent to the single equation

$$x+y=0.$$

Thus the eigenvectors corresponding to $\lambda=-8$ are the non-zero vectors (x,y) for which $y=-x$; that is, the vectors of the form

$$(k,-k),\quad \text{where } k\neq 0.$$

Thus the eigenvectors of t are the non-zero vectors of the following forms:

$$\begin{aligned}&(k,2k),\ \text{corresponding to } \lambda=1,\\ &(k,-k),\ \text{corresponding to } \lambda=-8.\end{aligned}$$

Solution to Exercise C119

(a) The matrix of t with respect to the standard basis for $\mathbb{R}^2$ is

$$\mathbf{A}=\begin{pmatrix}1 & 3\\ 2 & -4\end{pmatrix}.$$

We use Strategy C18 to find the eigenvalues and eigenvectors of $\mathbf{A}$, which are the same as those of t.

First we find the eigenvalues of $\mathbf{A}$.

The characteristic equation of $\mathbf{A}$ is $\det(\mathbf{A}-\lambda\mathbf{I})=0$; that is,

$$\begin{vmatrix}1-\lambda & 3\\ 2 & -4-\lambda\end{vmatrix}=0.$$

We expand the determinant and obtain

$$(1-\lambda)(-4-\lambda)-6=0,$$

which simplifies to

$$\lambda^2+3\lambda-10=(\lambda-2)(\lambda+5)=0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda=2$ and $\lambda=-5$.

Next we find the eigenvectors of $\mathbf{A}$.

The eigenvector equations are

$$\begin{aligned}(1-\lambda)x+{}&3y=0\\ 2x+(-4-\lambda)&y=0.\end{aligned}$$

$\boxed{\lambda=2}$ The eigenvector equations become

$$\begin{aligned}-x+3y&=0\\ 2x-6y&=0.\end{aligned}$$

These equations are equivalent to the single equation

$$x-3y=0.$$

Thus the eigenvectors corresponding to $\lambda=2$ are the non-zero vectors for which $x=3y$; that is, the vectors of the form

$$(3k,k),\quad \text{where } k\neq 0.$$

$\boxed{\lambda=-5}$ The eigenvector equations become

$$\begin{aligned}6x+3y&=0\\ 2x+\ \ y&=0.\end{aligned}$$

These equations are equivalent to the single equation

$$2x+y=0.$$

Thus the eigenvectors corresponding to $\lambda=-5$ are the non-zero vectors for which $y=-2x$; that is, the vectors of the form

$$(k,-2k),\quad \text{where } k\neq 0.$$

Thus the eigenvectors of t are the non-zero vectors of the following forms:

$$(3k, k),\ \text{corresponding to } \lambda = 2,$$
$$(k, -2k),\ \text{corresponding to } \lambda = -5.$$

(b) The matrix of t with respect to the standard basis for $\mathbb{R}^2$ is

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}.$$

We use Strategy C18 to find the eigenvalues and eigenvectors of $\mathbf{A}$, which are the same as those of t.

First we find the eigenvalues of $\mathbf{A}$.

The characteristic equation of $\mathbf{A}$ is $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$; that is,

$$\begin{vmatrix} 1-\lambda & -2 \\ -2 & -2-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(1-\lambda)(-2-\lambda) - 4 = 0,$$

which simplifies to

$$\lambda^2 + \lambda - 6 = (\lambda - 2)(\lambda + 3) = 0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda = 2$ and $\lambda = -3$.

Next we find the eigenvectors of $\mathbf{A}$.

The eigenvector equations are

$$\begin{aligned} (1-\lambda)x - \qquad 2y &= 0 \\ -2x + (-2-\lambda)y &= 0. \end{aligned}$$

$\boxed{\lambda = 2}$ The eigenvector equations become

$$\begin{aligned} -x - 2y &= 0 \\ -2x - 4y &= 0. \end{aligned}$$

These equations are equivalent to the single equation

$$x + 2y = 0.$$

Thus the eigenvectors corresponding to $\lambda = 2$ are the non-zero vectors for which $x = -2y$; that is, the vectors of the form

$$(-2k, k), \quad \text{where } k \neq 0.$$

$\boxed{\lambda = -3}$ The eigenvector equations become

$$\begin{aligned} 4x - 2y &= 0 \\ -2x + \ \ y &= 0. \end{aligned}$$

These equations are equivalent to the single equation

$$2x - y = 0.$$

Thus the eigenvectors corresponding to $\lambda = -3$ are the non-zero vectors for which $y = 2x$; that is, the vectors of the form

$$(k, 2k), \quad \text{where } k \neq 0.$$

Thus the eigenvectors of t are the non-zero vectors of the following forms:

$$(-2k, k),\ \text{corresponding to } \lambda = 2,$$
$$(k, 2k),\ \text{corresponding to } \lambda = -3.$$

Solution to Exercise C120

The matrix of t with respect to the standard basis for $\mathbb{R}^3$ is

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

We use Strategy C18 to find the eigenvalues and eigenvectors of $\mathbf{A}$, which are the same as those of t.

First we find the eigenvalues of $\mathbf{A}$.

The characteristic equation is $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$; that is,

$$\begin{vmatrix} 4-\lambda & 2 & 0 \\ 2 & 3-\lambda & 2 \\ 0 & 2 & 2-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(4-\lambda)\begin{vmatrix} 3-\lambda & 2 \\ 2 & 2-\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & 2 \\ 0 & 2-\lambda \end{vmatrix} + 0 = 0.$$

Simplifying this expression, we obtain

$$(4-\lambda)((3-\lambda)(2-\lambda) - 4) - 2(2(2-\lambda)) = 0,$$

or

$$\lambda^3 - 9\lambda^2 + 18\lambda = 0.$$

There is no constant term, so we take out the factor λ, then factorise the remaining quadratic factor:

$$\lambda(\lambda^2 - 9\lambda + 18) = \lambda(\lambda - 6)(\lambda - 3) = 0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda = 0$, $\lambda = 6$ and $\lambda = 3$.

(As a quick check $4+3+2=9=6+3+0$, so the sum of the eigenvalues is indeed equal to the sum of the diagonal entries.)

Next we find the eigenvectors of $\mathbf{A}$.

The eigenvector equations are

$$\begin{aligned}(4-\lambda)x + \quad 2y \qquad\quad &= 0\\ 2x + (3-\lambda)y + \quad 2z &= 0\\ 2y + (2-\lambda)z &= 0.\end{aligned}$$

$\boxed{\lambda = 6}$ The eigenvector equations become

$$\begin{aligned}-2x + 2y \qquad &= 0\\ 2x - 3y + 2z &= 0\\ 2y - 4z &= 0.\end{aligned}$$

The first and third equations imply that $x=y$ and $y=2z$, so $x=2z$. These satisfy the second equation. Thus the eigenvectors corresponding to the eigenvalue $\lambda=6$ are the non-zero vectors (x,y,z) satisfying $y=2z$ and $x=2z$; that is, the vectors of the form

$$(2k,2k,k), \quad \text{where } k\neq 0.$$

$\boxed{\lambda = 3}$ The eigenvector equations become

$$\begin{aligned}x + 2y \qquad &= 0\\ 2x \qquad + 2z &= 0\\ 2y - \quad z &= 0.\end{aligned}$$

The first and second equations imply that $x=-2y$ and $z=-x$, so $z=2y$. These satisfy the third equation. Thus the eigenvectors corresponding to the eigenvalue $\lambda=3$ are the non-zero vectors (x,y,z) satisfying $x=-2y$ and $z=2y$; that is, the vectors of the form

$$(-2k,k,2k), \quad \text{where } k\neq 0.$$

$\boxed{\lambda = 0}$ The eigenvector equations become

$$\begin{aligned}4x + 2y \qquad &= 0\\ 2x + 3y + 2z &= 0\\ 2y + 2z &= 0.\end{aligned}$$

The first and third equations imply that $y=-2x$ and $z=-y$, so $z=2x$. These satisfy the second equation. Thus the eigenvectors corresponding to the eigenvalue $\lambda=0$ are the non-zero vectors (x,y,z) satisfying $y=-2x$ and $z=2x$; that is, the vectors of the form

$$(k,-2k,2k), \quad \text{where } k\neq 0.$$

Thus the eigenvectors of t are the non-zero vectors of the following forms:

$$\begin{aligned}&(2k,2k,k), \text{ corresponding to } \lambda=6,\\ &(-2k,k,2k), \text{ corresponding to } \lambda=3,\\ &(k,-2k,2k), \text{ corresponding to } \lambda=0.\end{aligned}$$

Solution to Exercise C121

(a) Let

$$\mathbf{A} = \begin{pmatrix}1 & 2\\ 0 & 6\end{pmatrix}.$$

The characteristic equation is $\det(\mathbf{A}-\lambda\mathbf{I})=0$; that is,

$$\begin{vmatrix}1-\lambda & 2\\ 0 & 6-\lambda\end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(1-\lambda)(6-\lambda)-0=0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda=1$ and $\lambda=6$. Notice that these are the diagonal entries of the upper triangular matrix $\mathbf{A}$.

(b) Let

$$\mathbf{A} = \begin{pmatrix}8 & 0 & 0\\ 0 & -5 & 0\\ 0 & 0 & 21\end{pmatrix}.$$

The characteristic equation is $\det(\mathbf{A}-\lambda\mathbf{I})=0$; that is,

$$\begin{vmatrix}8-\lambda & 0 & 0\\ 0 & -5-\lambda & 0\\ 0 & 0 & 21-\lambda\end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(8-\lambda)\begin{vmatrix}-5-\lambda & 0\\ 0 & 21-\lambda\end{vmatrix} = 0.$$

Simplifying this expression, we obtain

$$(8-\lambda)((-5-\lambda)(21-\lambda)-0)=0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda=8$, $\lambda=-5$ and $\lambda=21$. Again, these are the diagonal entries of the diagonal matrix $\mathbf{A}$.

(c) Let

$$\mathbf{A} = \begin{pmatrix} 4 & 0 & 0 \\ 25 & -2 & 0 \\ 17 & \pi & 6 \end{pmatrix}.$$

The characteristic equation is $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$; that is,

$$\begin{vmatrix} 4-\lambda & 0 & 0 \\ 25 & -2-\lambda & 0 \\ 17 & \pi & 6-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(4-\lambda)\begin{vmatrix} -2-\lambda & 0 \\ \pi & 6-\lambda \end{vmatrix} = 0.$$

Simplifying this expression, we obtain

$$(4-\lambda)((-2-\lambda)(6-\lambda) - 0) = 0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda = 4$, $\lambda = -2$ and $\lambda = 6$. Again, these are the diagonal entries of the lower triangular matrix $\mathbf{A}$.

Solution to Exercise C122

$\boxed{\lambda = 6}$ The non-zero vectors of the form $(2k, 2k, k)$ are the eigenvectors of t corresponding to $\lambda = 6$. The eigenspace $S(6)$ is therefore the set of vectors

$$\{(2k, 2k, k) : k \in \mathbb{R}\}.$$

Any vector in $S(6)$ can be written as $k(2, 2, 1)$, so $\{(2, 2, 1)\}$ is a basis for $S(6)$.

Thus $S(6)$ has dimension 1.

$\boxed{\lambda = 3}$ The non-zero vectors of the form $(-2k, k, 2k)$ are the eigenvectors of t corresponding to $\lambda = 3$. The eigenspace $S(3)$ is therefore the set of vectors

$$\{(-2k, k, 2k) : k \in \mathbb{R}\}.$$

Any vector in $S(3)$ can be written as $k(-2, 1, 2)$, so $\{(-2, 1, 2)\}$ is a basis for $S(3)$.

Thus $S(3)$ has dimension 1.

Solution to Exercise C123

The matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

is triangular, so the eigenvalues are the diagonal entries $\lambda = 1$, $\lambda = 4$ and $\lambda = 4$.

The eigenvector equations are

$$\begin{aligned} (1-\lambda)x + \quad y - \quad z &= 0 \\ (4-\lambda)y \quad &= 0 \\ (4-\lambda)z &= 0. \end{aligned}$$

$\boxed{\lambda = 1}$ The eigenvalue $\lambda = 1$ has multiplicity 1.

The eigenvector equations become

$$\begin{aligned} y - z &= 0 \\ 3y \quad &= 0 \\ 3z &= 0. \end{aligned}$$

The second and third equations give $y = 0$ and $z = 0$, respectively, which satisfy the first equation. (They give no constraint on x.)

Thus the eigenvectors corresponding to the eigenvalue $\lambda = 1$ are the vectors of the form $(k, 0, 0)$, where $k \neq 0$.

The eigenspace $S(1)$ is the set of vectors

$$\{(k, 0, 0) : k \in \mathbb{R}\}.$$

Any vector in $S(1)$ can be written as $k(1, 0, 0)$, so

$$\{(1, 0, 0)\}$$

is a basis for $S(1)$.

Thus $S(1)$ has dimension 1.

(Geometrically, $S(1)$ is the x-axis.)

$\boxed{\lambda = 4}$ The eigenvalue $\lambda = 4$ has multiplicity 2.

The eigenvector equations become

$$\begin{aligned} -3x + y - z &= 0 \\ 0y \quad &= 0 \\ 0z &= 0. \end{aligned}$$

The first equation gives $z = y - 3x$ and the second and third give no constraints on y and z.

Thus the eigenvectors corresponding to the eigenvalue $\lambda = 4$ are the vectors of the form $(k, l, l - 3k)$, where k and l are not both 0.

The eigenspace $S(4)$ is the set of vectors

$$\{(k, l, l - 3k) : k, l \in \mathbb{R}\}.$$

Any vector in $S(4)$ can be written as $k(1,0,-3)+l(0,1,1)$, so

$$\{(1,0,-3),(0,1,1)\}$$

is a basis for $S(4)$.

Thus $S(4)$ has dimension 2.

(Geometrically, $S(4)$ is the plane in $\mathbb{R}^3$ $-3x+y-z=0$.)

An alternative solution comes from using the equivalent equation $x=\frac{1}{3}(y-z)$, and has basis

$$\left\{\left(\tfrac{1}{3},1,0\right),\left(-\tfrac{1}{3},0,1\right)\right\}.$$

Solution to Exercise C124

The matrix

$$\mathbf{A}=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$$

is triangular, so the eigenvalues are the diagonal entries $\lambda=1$ and $\lambda=1$.

The eigenvector equations are

$$\begin{aligned}(1-\lambda)x + \quad\;\; y &= 0\\ (1-\lambda)y &= 0.\end{aligned}$$

$\boxed{\lambda=1}$ The eigenvalue $\lambda=1$ has multiplicity 2.

The eigenvector equations become

$$\begin{aligned}0x + \;\; y &= 0\\ 0y &= 0.\end{aligned}$$

Thus $y=0$ and there are no constraints on x. Thus the eigenvectors corresponding to the eigenvalue $\lambda=1$ are the vectors of the form $(k,0)$, where $k\neq 0$.

The eigenspace $S(1)$ is the set of vectors

$$\{(k,0):k\in\mathbb{R}\}.$$

Any vector in $S(1)$ can be written as $k(1,0)$, so

$$\{(1,0)\}$$

is a basis for $S(1)$.

Thus $S(1)$ has dimension 1.

(Geometrically, $S(1)$ is the x-axis in $\mathbb{R}^2$.)

Solution to Exercise C125

Let

$$\mathbf{A}=\begin{pmatrix}1 & -1 & 0\\ 1 & 4 & 1\\ -1 & 1 & 4\end{pmatrix}.$$

The characteristic equation is $\det(\mathbf{A}-\lambda\mathbf{I})=0$; that is,

$$\begin{vmatrix}1-\lambda & -1 & 0\\ 1 & 4-\lambda & 1\\ -1 & 1 & 4-\lambda\end{vmatrix}=0.$$

We expand the determinant and obtain

$$(1-\lambda)\begin{vmatrix}4-\lambda & 1\\ 1 & 4-\lambda\end{vmatrix}+\begin{vmatrix}1 & 1\\ -1 & 4-\lambda\end{vmatrix}=0.$$

This simplifies to

$$(1-\lambda)((4-\lambda)^2-1)+((4-\lambda)+1)=0.$$

Using the relation $x^2-1=(x-1)(x+1)$, where $x=4-\lambda$, this simplifies further to

$$(1-\lambda)(3-\lambda)(5-\lambda)+(5-\lambda)=0,$$

and thus

$$\begin{aligned}(5-\lambda)((1-\lambda)(3-\lambda)+1)&=(5-\lambda)(\lambda^2-4\lambda+4)\\ &=(5-\lambda)(\lambda-2)^2\\ &=0.\end{aligned}$$

The eigenvalues of $\mathbf{A}$ are $\lambda=5$, $\lambda=2$ and $\lambda=2$.

(As a quick check $1+4+4=9=5+2+2$, so the sum of the eigenvalues is indeed equal to the sum of the diagonal entries.)

The eigenvector equations are

$$\begin{aligned}(1-\lambda)x - \qquad y \qquad\qquad &= 0\\ x + (4-\lambda)y + \qquad z &= 0\\ -x + \qquad y + (4-\lambda)z &= 0.\end{aligned}$$

$\boxed{\lambda=5}$ The eigenvalue $\lambda=5$ has multiplicity 1.

The eigenvector equations become

$$\begin{aligned}-4x - y \qquad &= 0\\ x - y + z &= 0\\ -x + y - z &= 0.\end{aligned}$$

The first equation gives $y=-4x$ and substituting this into the second gives $5x+z=0$, which implies that $z=-5x$. The third equation is equivalent to the second.

Thus the eigenvectors corresponding to the eigenvalue $\lambda = 5$ are the vectors of the form $(k, -4k, -5k)$, where $k \neq 0$.

The eigenspace $S(5)$ is the set of vectors

$$\{(k, -4k, -5k) : k \in \mathbb{R}\}.$$

Any vector in $S(5)$ can be written as $k(1, -4, -5)$, so

$$\{(1, -4, -5)\}$$

is a basis for $S(5)$.

Thus $S(5)$ has dimension 1.

$\boxed{\lambda = 2}$ The eigenvalue $\lambda = 2$ has multiplicity 2.

The eigenvector equations become

$$\begin{aligned} -x - y \phantom{{}+2z} &= 0 \\ x + 2y + z &= 0 \\ -x + y + 2z &= 0. \end{aligned}$$

The first equation gives $y = -x$ and substituting this into the second gives $-x + z = 0$, which implies that $z = x$. These satisfy the third equation.

Thus the eigenvectors corresponding to the eigenvalue $\lambda = 2$ are the vectors of the form $(k, -k, k)$, where $k \neq 0$.

The eigenspace $S(2)$ is the set of vectors

$$\{(k, -k, k) : k \in \mathbb{R}\}.$$

Any vector in $S(2)$ can be written as $k(1, -1, 1)$, so

$$\{(1, -1, 1)\}$$

is a basis for $S(2)$.

Thus $S(2)$ has dimension 1.

Solution to Exercise C126

Letting $k = 1$, we see that $(-2, 1)$ and $(1, 2)$ are eigenvectors of t. Since $(1, 2)$ is not a multiple of $(-2, 1)$, these two eigenvectors form a basis for $\mathbb{R}^2$.

Solution to Exercise C127

Each of the vectors in E is an eigenvector of t:

$$\begin{aligned} t(0, 1, -1) &= (0, 0, 0) = 0(0, 1, -1), \\ t(-2, 1, 0) &= (4, -2, 0) = -2(-2, 1, 0), \\ t(1, 0, -1) &= (-3, 0, 3) = -3(1, 0, -1). \end{aligned}$$

Thus E is a basis for $\mathbb{R}^3$ consisting of eigenvectors of t; that is, E is an eigenvector basis of t.

Solution to Exercise C128

(a) The matrix of t with respect to the standard basis for $\mathbb{R}^2$ is

$$\begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}.$$

(b) Following Strategy C19, first we find the images of the vectors in the basis $E = \{(-2, 1), (1, 2)\}$:

$$t(-2, 1) = (-4, 2), \quad t(1, 2) = (-3, -6).$$

Next we find the E-coordinates of each of these image vectors:

$$\begin{aligned} (-4, 2) &= 2(-2, 1) + 0(1, 2) \\ &= (2, 0)_E, \\ (-3, -6) &= 0(-2, 1) - 3(1, 2) \\ &= (0, -3)_E. \end{aligned}$$

Therefore $t(-2, 1) = (2, 0)_E$ and $t(1, 2) = (0, -3)_E$. So the matrix of t with respect to the eigenvector basis E is

$$\begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}.$$

Solution to Exercise C129

In Exercise C127 you showed that

$$\begin{aligned} t(0, 1, -1) &= 0(0, 1, -1), \\ t(-2, 1, 0) &= -2(-2, 1, 0), \\ t(1, 0, -1) &= -3(1, 0, -1). \end{aligned}$$

So the eigenvalues of t are $\lambda_1 = 0$, $\lambda_2 = -2$ and $\lambda_3 = -3$, and, by Theorem C59, the matrix of t with respect to E is

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}.$$

Solution to Exercise C130

(a) $\mathbf{P} = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$

(b) $\mathbf{P} = \begin{pmatrix} 0 & -2 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$

Solution to Exercise C131

Let $t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the linear transformation given by

$$t(x, y) = (x - 2y, -2x - 2y)$$

and let E be the eigenvector basis $\{(-2, 1), (1, 2)\}$ of t. It follows from Exercise C128 that $\mathbf{A}$ is the matrix of t with respect to the standard basis for $\mathbb{R}^2$ and $\mathbf{D}$ is the matrix of t with respect to the eigenvector basis E. By Theorem C62, $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$, where $\mathbf{P}$ is the transition matrix from E to the standard basis for $\mathbb{R}^2$; that is,

$$\mathbf{P} = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Solution to Exercise C132

(a) $\mathbf{D}^5 = \begin{pmatrix} 2^5 & 0 \\ 0 & (-3)^5 \end{pmatrix} = \begin{pmatrix} 32 & 0 \\ 0 & -243 \end{pmatrix}$

(b) We have $\mathbf{A}^5 = \mathbf{PD}^5\mathbf{P}^{-1}$, where $\mathbf{D}$ is as in part (a) and

$$\mathbf{P} = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Since $\mathbf{P}^{-1} = -\frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & -2 \end{pmatrix}$, it follows that

$$\begin{aligned}\mathbf{A}^5 &= \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 32 & 0 \\ 0 & -243 \end{pmatrix}\begin{pmatrix} -\frac{2}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix} \\ &= \begin{pmatrix} -23 & -110 \\ -110 & -188 \end{pmatrix}.\end{aligned}$$

Solution to Exercise C133

(There are many solutions possible for this and for each of the remaining exercises in this section, each corresponding to a different ordering of the eigenvalues or a different choice of eigenvectors; in each case the matrix $\mathbf{P}$ should correspond to the matrix $\mathbf{D}$ so that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$.)

We use Strategy C20.

The eigenvalues of $\mathbf{A}$ are $\lambda = 6$, $\lambda = 3$ and $\lambda = 0$.

The eigenvectors of $\mathbf{A}$ are the non-zero vectors of the following forms:

$$\begin{aligned}&(2k, 2k, k), \text{ corresponding to } \lambda = 6,\\ &(-2k, k, 2k), \text{ corresponding to } \lambda = 3,\\ &(k, -2k, 2k), \text{ corresponding to } \lambda = 0.\end{aligned}$$

It follows from Theorem C63 that we can form an eigenvector basis of $\mathbf{A}$ by taking one eigenvector corresponding to each of the three distinct eigenvalues. For example,

$$E = \{(2, 2, 1), (-2, 1, 2), (1, -2, 2)\}$$

is an eigenvector basis of $\mathbf{A}$.

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{pmatrix}.$$

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Solution to Exercise C134

The characteristic equation of $\mathbf{A}$ is

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(1-\lambda)((2-\lambda)^2 - 1) = 0,$$

which simplifies to

$$\begin{aligned}(1-\lambda)(\lambda^2 - 4\lambda + 3) &= (1-\lambda)(\lambda - 1)(\lambda - 3) \\ &= 0.\end{aligned}$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda = 3$, $\lambda = 1$ and $\lambda = 1$.

To find the eigenspaces of $\mathbf{A}$, we consider the eigenvector equations

$$\begin{aligned}(1-\lambda)x \qquad\qquad\qquad\qquad &= 0\\ (2-\lambda)y + \qquad\quad z &= 0\\ y + (2-\lambda)z &= 0,\end{aligned}$$

for each of the eigenvalues.

$\boxed{\lambda = 3}$ The eigenvector equations become

$$\begin{aligned} -2x \qquad\quad &= 0 \\ -y + z &= 0 \\ y - z &= 0. \end{aligned}$$

So $x = 0$, $y = z$.

Thus $S(3) = \{(0, k, k) : k \in \mathbb{R}\}$.

$\boxed{\lambda = 1}$ The eigenvector equations become

$$\begin{aligned} 0x \qquad\quad &= 0 \\ y + z &= 0 \\ y + z &= 0. \end{aligned}$$

So $z = -y$ and there are no constraints on x.

Thus $S(1) = \{(k, l, -l) : k, l \in \mathbb{R}\}$.

A basis for $S(3)$ is $\{(0, 1, 1)\}$ and a basis for $S(1)$ is $\{(1, 0, 0), (0, 1, -1)\}$ because any vector in $S(1)$ can be written as $k(1, 0, 0) + l(0, 1, -1)$. The set

$$E = \{(0, 1, 1), (1, 0, 0), (0, 1, -1)\}$$

contains three vectors, so it is an eigenvector basis of $\mathbf{A}$.

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solution to Exercise C135

(a) We have

$$\begin{aligned} (2k, 2k, k) \cdot (-2l, l, 2l) &= -4kl + 2kl + 2kl \\ &= 0, \\ (2k, 2k, k) \cdot (m, -2m, 2m) &= 2km - 4km + 2km \\ &= 0, \\ (-2l, l, 2l) \cdot (m, -2m, 2m) &= -2lm - 2lm + 4lm \\ &= 0. \end{aligned}$$

Thus the given vectors form an orthogonal set. Since there are three of them, they form an orthogonal basis for $\mathbb{R}^3$.

(b)

$$\begin{aligned} |\mathbf{v}_1| = |(2k, 2k, k)| &= \sqrt{4k^2 + 4k^2 + k^2} \\ &= \sqrt{9k^2} \\ &= 3k, \\ |\mathbf{v}_2| = |(-2l, l, 2l)| &= \sqrt{4l^2 + l^2 + 4l^2} \\ &= \sqrt{9l^2} \\ &= 3l, \\ |\mathbf{v}_3| = |(m, -2m, 2m)| &= \sqrt{m^2 + 4m^2 + 4m^2} \\ &= \sqrt{9m^2} \\ &= 3m. \end{aligned}$$

Thus $|\mathbf{v}_1| = |\mathbf{v}_2| = |\mathbf{v}_3| = 1$ if

$$k = l = m = \tfrac{1}{3}.$$

Solution to Exercise C136

We calculate $\mathbf{P}^T\mathbf{P}$.

$$\begin{aligned} \mathbf{P}^T\mathbf{P} &= \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix} \\ &= \begin{pmatrix} \frac{9}{9} & 0 & 0 \\ 0 & \frac{9}{9} & 0 \\ 0 & 0 & \frac{9}{9} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I} \end{aligned}$$

Solution to Exercise C137

(a) We use Strategy C22.

The characteristic equation of $\mathbf{A}$ is

$$\begin{vmatrix} 9 - \lambda & -2 \\ -2 & 6 - \lambda \end{vmatrix} = 0.$$

We expand the determinant and obtain

$$(9 - \lambda)(6 - \lambda) - 4 = 0,$$

which simplifies to

$$\lambda^2 - 15\lambda + 50 = (\lambda - 10)(\lambda - 5) = 0.$$

The eigenvalues of $\mathbf{A}$ are therefore $\lambda = 10$ and $\lambda = 5$.

Next we find orthonormal bases for the eigenspaces.

The eigenvector equations are

$$\begin{aligned} (9 - \lambda)x - \qquad 2y &= 0 \\ -2x + (6 - \lambda)y &= 0. \end{aligned}$$

$\boxed{\lambda = 10}$ The eigenvector equations become

$$\begin{aligned} -x - 2y &= 0 \\ -2x - 4y &= 0. \end{aligned}$$

These equations are equivalent to the single equation

$$x + 2y = 0,$$

that is, $x = -2y$. Thus the eigenvectors corresponding to $\lambda = 10$ are the non-zero vectors of the form $(-2k, k)$.

An eigenvector of magnitude 1 corresponding to $\lambda = 10$ is

$$\left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right).$$

$\boxed{\lambda = 5}$ The eigenvector equations become

$$\begin{aligned} 4x - 2y &= 0 \\ -2x + y &= 0. \end{aligned}$$

These equations are equivalent to the single equation

$$2x - y = 0,$$

that is, $y = 2x$. Thus the eigenvectors corresponding to $\lambda = 5$ are the non-zero vectors of the form $(k, 2k)$.

An eigenvector of magnitude 1 corresponding to $\lambda = 5$ is

$$\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right).$$

It follows from Theorem C64 that an orthonormal eigenvector basis of $\mathbf{A}$ is

$$E = \left\{\left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right), \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)\right\}.$$

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}.$$

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}.$$

(b) The eigenvalues of $\mathbf{A}$ are given as $\lambda = 6$, $\lambda = 3$ and $\lambda = 2$.

Now we find an orthonormal eigenvector basis of $\mathbf{A}$.

The eigenvector equations are

$$\begin{aligned} (5-\lambda)x - y - z &= 0 \\ -x + (3-\lambda)y + z &= 0 \\ -x + y + (3-\lambda)z &= 0. \end{aligned}$$

$\boxed{\lambda = 6}$ The eigenvector equations become

$$\begin{aligned} -x - y - z &= 0 \\ -x - 3y + z &= 0 \\ -x + y - 3z &= 0. \end{aligned}$$

Adding the first and second equations together, we obtain

$$-2x - 4y = 0,$$

so $x = -2y$. Substituting this into the third equation, we obtain

$$3y - 3z = 0,$$

so $z = y$. Thus the eigenvectors corresponding to $\lambda = 6$ are the non-zero vectors of the form $(-2k, k, k)$.

An eigenvector of magnitude 1 corresponding to $\lambda = 6$ is

$$\left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right).$$

$\boxed{\lambda = 3}$ The eigenvector equations become

$$\begin{aligned} 2x - y - z &= 0 \\ -x \phantom{{}-y} + z &= 0 \\ -x + y \phantom{{}-z} &= 0. \end{aligned}$$

The second and third equations imply that $z = x$ and $y = x$. These satisfy the first equation. Thus the eigenvectors corresponding to $\lambda = 3$ are the non-zero vectors of the form (k, k, k).

An eigenvector of magnitude 1 corresponding to $\lambda = 3$ is

$$\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

$\boxed{\lambda = 2}$ The eigenvector equations become

$$\begin{aligned} 3x - y - z &= 0 \\ -x + y + z &= 0 \\ -x + y + z &= 0. \end{aligned}$$

Adding the first and second equations together, we obtain

$$2x = 0,$$

which implies that $x = 0$. Substituting this into the third equation, we obtain

$$y + z = 0,$$

which implies that $z = -y$. Thus the eigenvectors corresponding to $\lambda = 2$ are the non-zero vectors of the form $(0, k, -k)$.

An eigenvector of magnitude 1 corresponding to $\lambda = 2$ is

$$\left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

It follows from Theorem C64 that an orthonormal eigenvector basis of $\mathbf{A}$ is

$$E = \left\{\left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right), \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\right\}.$$

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Solution to Exercise C138

We use Strategy C22.

A basis for the eigenspace $S(3)$ is $\{(0, 1, 1)\}$, so an orthonormal basis for $S(3)$ is

$$\left\{\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right\}.$$

A basis for the eigenspace $S(1)$ is $\{(1, 0, 0), (0, 1, -1)\}$.

These two basis vectors are orthogonal since

$$(1, 0, 0) \cdot (0, 1, -1) = 0.$$

An orthonormal basis for $S(1)$ is therefore

$$\left\{(1, 0, 0), \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\right\}.$$

By Theorem C64 an orthonormal eigenvector basis of $\mathbf{A}$ is therefore

$$E = \left\{\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), (1, 0, 0), \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\right\}.$$

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

We use the eigenvalues corresponding to the eigenvectors in E to form the diagonal matrix:

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solution to Exercise C139

By Theorem C65, to prove that the product $\mathbf{PQ}$ is orthogonal it is sufficient to show that

$$(\mathbf{PQ})^T = (\mathbf{PQ})^{-1}.$$

But

$$(\mathbf{PQ})^T = \mathbf{Q}^T\mathbf{P}^T = \mathbf{Q}^{-1}\mathbf{P}^{-1} = (\mathbf{PQ})^{-1}.$$

Solution to Exercise C140

(a) To verify that $\mathbf{A}$ is orthogonal, it is sufficient to show that $\mathbf{A}^T\mathbf{A} = \mathbf{I}$, by Theorem C65.

$$\begin{aligned} \mathbf{A}^T\mathbf{A} &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I}, \end{aligned}$$

so $\mathbf{A}$ is orthogonal.

(Alternatively, we could have shown that the vectors $(0, 0, 1)$, $(0, 1, 0)$ and $(-1, 0, 0)$ form an orthonormal basis for $\mathbb{R}^3$.)

(b) We evaluate the determinant of $\mathbf{A}$:

$$\begin{vmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 0 - 0 - \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 1.$$

Therefore $\mathbf{A}$ represents a rotation of $\mathbb{R}^3$.

Solution to Exercise C141

(a) The ellipse with equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

is written in matrix form as

$$\mathbf{x}^T \begin{pmatrix} 1/a^2 & 0 \\ 0 & 1/b^2 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 & 0 \end{pmatrix} \mathbf{x} - 1 = 0.$$

So the ellipse in standard position has

$$\mathbf{A} = \begin{pmatrix} 1/a^2 & 0 \\ 0 & 1/b^2 \end{pmatrix} \text{ and } \mathbf{J} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

(b) The hyperbola with equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

is written in matrix form as

$$\mathbf{x}^T \begin{pmatrix} 1/a^2 & 0 \\ 0 & -1/b^2 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 & 0 \end{pmatrix} \mathbf{x} - 1 = 0.$$

So the hyperbola in standard position has

$$\mathbf{A} = \begin{pmatrix} 1/a^2 & 0 \\ 0 & -1/b^2 \end{pmatrix} \text{ and } \mathbf{J} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

(c) The parabola with equation

$$y^2 = 4ax$$

is written in matrix form as

$$\mathbf{x}^T \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} -4a & 0 \end{pmatrix} \mathbf{x} + 0 = 0.$$

So the parabola in standard position has

$$\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \mathbf{J} = \begin{pmatrix} -4a \\ 0 \end{pmatrix}.$$

Solution to Exercise C142

By Theorem C64 an orthonormal eigenvector basis of $\mathbf{A}$ for the eigenvalues $\lambda = 2$ and $\lambda = -3$, in that order, is

$$E = \left\{ \left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right), \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right) \right\}.$$

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}.$$

We use the eigenvalues to form the diagonal matrix

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{D} = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}.$$

It follows from equation (9) that the equation of the conic is now

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 6 & 12 \end{pmatrix} \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + 21 = 0,$$

that is,

$$2(x')^2 - 3(y')^2 + 6\sqrt{5}y' + 21 = 0.$$

Solution to Exercise C143

We have

$$2(x')^2 - 3(y')^2 + 6\sqrt{5}y' + 21 = 0,$$

which is equivalent to

$$2(x')^2 - 3\left((y')^2 - 2\sqrt{5}y'\right) + 21 = 0.$$

Completing the square gives

$$2(x')^2 - 3(y' - \sqrt{5})^2 + 15 + 21 = 0,$$

so

$$2(x')^2 - 3(y' - \sqrt{5})^2 + 36 = 0.$$

We set the new coordinates to be

$$\mathbf{x}'' = (x'', y'') = (x', y' - \sqrt{5}),$$

so substitute $x'' = x'$ and $y'' = y' - \sqrt{5}$.

The equation of the conic is now

$$2(x'')^2 - 3(y'')^2 = -36,$$

or

$$-\frac{(x'')^2}{18} + \frac{(y'')^2}{12} = 1.$$

Solution to Exercise C144

1. Introduce matrices.

   We have

   $$\mathbf{A} = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix} \quad \text{and} \quad \mathbf{J} = \begin{pmatrix} -10 \\ -20 \end{pmatrix}.$$

2. Align the axes.

   We have

   $$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix},$$

   where

   $$\mathbf{P} = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}.$$

   So

   $$\begin{aligned}(f \quad g) &= (-10 \quad -20)\begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \\ &= \left(0 \quad -\frac{50}{\sqrt{5}}\right) \\ &= \left(0 \quad -10\sqrt{5}\right).\end{aligned}$$

   The equation of the conic is now

   $$10(x')^2 + 5(y')^2 - 10\sqrt{5}y' - 5 = 0.$$

   Dividing through by 5, we obtain

   $$2(x')^2 + (y')^2 - 2\sqrt{5}y' - 1 = 0.$$

3. Translate the origin.

   We write this equation as

   $$2(x')^2 + \left((y')^2 - 2\sqrt{5}y'\right) - 1 = 0.$$

   Completing the square in this equation, we obtain

   $$2(x')^2 + (y' - \sqrt{5})^2 - 5 - 1 = 0.$$

   Substituting $x'' = x'$ and $y'' = y' - \sqrt{5}$ in this equation and simplifying, we obtain

   $$2(x'')^2 + (y'')^2 - 6 = 0.$$

The equation of the conic in standard form is

$$\frac{(x'')^2}{3} + \frac{(y'')^2}{6} = 1.$$

The conic is an ellipse.

Solution to Exercise C145

1. Introduce matrices.

   We have

   $$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{J} = \begin{pmatrix} -6 \\ -8 \end{pmatrix}.$$

2. Align the axes.

   The characteristic equation of $\mathbf{A}$ is

   $$\begin{vmatrix} 1-\lambda & -2 \\ -2 & 4-\lambda \end{vmatrix} = 0.$$

   We expand the determinant and obtain

   $$(1-\lambda)(4-\lambda) - 4 = 0,$$

   which simplifies to

   $$\lambda^2 - 5\lambda = \lambda(\lambda - 5) = 0.$$

   The eigenvalues of $\mathbf{A}$ are 5 and 0.

   The eigenvector equations are

   $$\begin{aligned}(1-\lambda)x - \quad\quad 2y &= 0 \\ -2x + (4-\lambda)y &= 0.\end{aligned}$$

   $\boxed{\lambda = 5}$ The eigenvector equations become

   $$\begin{aligned}-4x - 2y &= 0 \\ -2x - y &= 0.\end{aligned}$$

   These equations are equivalent to the single equation

   $$2x + y = 0,$$

   which implies that $y = -2x$. Thus the eigenvectors corresponding to $\lambda = 5$ are the non-zero vectors of the form $(k, -2k)$.

   An orthonormal basis for $S(5)$ is

   $$\left\{\left(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right)\right\}.$$

   $\boxed{\lambda = 0}$ The eigenvector equations become

   $$\begin{aligned}x - 2y &= 0 \\ -2x + 4y &= 0.\end{aligned}$$

   These equations are equivalent to the single equation

   $$x - 2y = 0,$$

   which implies that $x = 2y$. Thus the eigenvectors corresponding to $\lambda = 0$ are the non-zero vectors of the form $(2k, k)$.

An orthonormal basis for $S(0)$ is

$$\left\{\left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)\right\}.$$

By Theorem C64 an orthonormal eigenvector basis of $\mathbf{A}$ is therefore

$$E = \left\{\left(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right), \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)\right\}.$$

We use the eigenvectors in E to form the columns of the transition matrix:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}.$$

Now,

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}$$

so

$$\begin{aligned}(f \quad g) &= (-6 \quad -8)\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{10}{\sqrt{5}} & -\frac{20}{\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} 2\sqrt{5} & -4\sqrt{5} \end{pmatrix}.\end{aligned}$$

The equation of the conic is now

$$5(x')^2 + 2\sqrt{5}x' - 4\sqrt{5}y' + 5 = 0.$$

3. Translate the origin.

We rewrite this equation by taking out the coefficient of the $(x')^2$ term to get

$$5\left((x')^2 + \frac{2}{\sqrt{5}}x'\right) - 4\sqrt{5}y' + 5 = 0.$$

Completing the square in this equation, we obtain

$$5\left(x' + \frac{1}{\sqrt{5}}\right)^2 - 1 - 4\sqrt{5}y' + 5 = 0.$$

We substitute

$$x'' = x' + \frac{1}{\sqrt{5}}$$

into this equation and rewrite it by taking out the coefficient of the y' term to get

$$5(x'')^2 - 4\sqrt{5}\left(y' - \frac{1}{\sqrt{5}}\right) = 0.$$

We substitute

$$y'' = y' - \frac{1}{\sqrt{5}}$$

to obtain

$$5(x'')^2 - 4\sqrt{5}y'' = 0.$$

The equation of the conic in standard form is

$$(x'')^2 = \frac{4}{\sqrt{5}}y''.$$

The conic is a parabola.

Solution to Exercise C146

1. Introduce matrices.

We have

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} -2 \\ 4 \\ -6 \end{pmatrix}.$$

2. Align the axes.

The matrix is already in diagonal form. (The axes of the quadric are parallel to the x-axis, y-axis and z-axis of $\mathbb{R}^3$.)

3. Translate the origin.

We write the equation as

$$(x^2 - 2x) + (y^2 + 4y) + (z^2 - 6z) - 11 = 0.$$

Completing the squares in this equation, we obtain

$$\begin{aligned}(x-1)^2 - 1 + (y+2)^2 - 4 \\ + (z-3)^2 - 9 - 11 = 0.\end{aligned}$$

Substituting

$$x' = x - 1, \quad y' = y + 2 \quad \text{and} \quad z' = z - 3$$

in this equation and simplifying, we obtain

$$(x')^2 + (y')^2 + (z')^2 - 25 = 0.$$

The equation of the quadric in standard form is

$$\frac{(x')^2}{25} + \frac{(y')^2}{25} + \frac{(z')^2}{25} = 1.$$

This is the equation of an ellipsoid.

(This ellipsoid is in fact a sphere since $a = b = c = 5$; all the curves of intersection are circles.)

Solution to Exercise C147

1. Introduce matrices. We have

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 12 \\ 0 \\ 12 \end{pmatrix}.$$

2. Align the axes.

   We have

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

   where

$$\mathbf{P} = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}.$$

   So

$$\begin{aligned} \begin{pmatrix} f & g & h \end{pmatrix} &= \begin{pmatrix} 12 & 0 & 12 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix} \\ &= \begin{pmatrix} 12 & 0 & 12 \end{pmatrix}. \end{aligned}$$

   The equation of the quadric is now

$$6(x')^2 + 3(y')^2 + 12x' + 12z' + 18 = 0.$$

3. Translate the origin.

   We write this equation as

$$6\left((x')^2 + 2x'\right) + 3(y')^2 + 12z' + 18 = 0.$$

   Completing the square in this equation, we obtain

$$6(x'+1)^2 - 6 + 3(y')^2 + 12z' + 18 = 0.$$

   Substituting

$$x'' = x' + 1, \quad y'' = y' \quad \text{and} \quad z'' = z' + 1$$

   in this equation and simplifying, we obtain

$$2(x'')^2 + (y'')^2 + 4z'' = 0.$$

The equation of the quadric in standard form is

$$\frac{(x'')^2}{2} + \frac{(y'')^2}{4} = -z''.$$

This is the equation of an elliptic paraboloid.